```
                    GenCore version 5.1.6
          Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:        January 8, 2006, 15:39:24 ; Search time 101 Seconds
                                          (without alignments)
                                          527.989 Million cell updates/sec

Title:          US-10-791-367-1
Perfect score:  30
Sequence:       1 caggcuacgcgtagagcaucatgatccugt 30

Scoring table:  IDENTITY.NUC
                Gapop 10.0 , Gapext 1.0

Searched:       130057 seqs, 88878828 residues

Total number of hits satisfying chosen parameters:   2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :      Issued Patents NA:*
                1: /cgn2_6/ptodata/1/ina/1_COMB.seq:*
                2: /cgn2_6/ptodata/1/ina/5_COMB.seq:*
                3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
                4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
                5: /cgn2_6/ptodata/1/ina/H_COMB.seq:*
                6: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
                7: /cgn2_6/ptodata/1/ina/PP_COMB.seq:*
                8: /cgn2_6/ptodata/1/ina/RE_COMB.seq:*
                9: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	30	100.0	30	3	US-09-859-724-1	Sequence 1, Appli
2	30	100.0	31	3	US-09-246-461-3	Sequence 3, Appli
3	30	100.0	31	3	US-10-021-330-3	Sequence 3, Appli
4	19.6	65.3	606	3	US-09-270-767-11347	Sequence 11347, A
5	19.6	65.3	615	3	US-09-270-767-14489	Sequence 14489, A
6	19.6	65.3	705	3	US-09-270-767-10626	Sequence 10626, A
7	19.6	65.3	167708	3	US-09-949-016-16423	Sequence 16423, A
8	.19	63.3	167708	3	US-09-489-039A-2715	Sequence 2715, Ap
9	18.8	62.7	1920	3	US-09-447-497-4	Sequence 4, Appli
10	18.8	62.7	1932	3	US-09-447-497-1	Sequence 1, Appli
11	18.4	61.3	485	3	US-09-533-559-2284	Sequence 2284, Ap
c 12	18	60.0	50	3	US-08-753-247-7	Sequence 7, Appli
13	18	60.0	477	3	US-09-270-767-8822	Sequence 8822, Ap
14	18	60.0	477	3	US-09-270-767-24104	Sequence 24104, A
15	18	60.0	2142	3	US-08-753-247-8	Sequence 8, Appli
16	17.8	59.3	601	3	US-09-949-016-44850	Sequence 44850, A
17	17.8	59.3	601	3	US-09-949-016-44851	Sequence 44851, A
18	17.8	59.3	601	3	US-09-949-016-73896	Sequence 73896, A
19	17.8	59.3	601	3	US-09-949-016-193773	Sequence 193773,
c 20	17.8	59.3	601	3	US-09-949-002-4223	Sequence 4223, Ap
c 21	17.8	59.3	43690	3	US-09-949-016-13904	Sequence 13904, A
c 22	17.8	59.3	67745	3	US-09-949-016-17251	Sequence 17251, A
c 23	17.8	59.3	105189	3	US-09-949-016-13029	Sequence 13029, A
c 24	17.8	59.3	129554	3	US-09-949-002-765	Sequence 765, App
25	17.8	59.3	179904	3	US-09-949-002-577	Sequence 577, App
26	17.8	59.3	179905	3	US-09-949-002-705	Sequence 705, App
27	17.6	58.7	3375	2	US-08-381-931B-1	Sequence 1, Appli
28	17.4	58.0	246	3	US-09-388-552-59	Sequence 59, Appli
c 29	17.4	58.0	705	3	US-08-781-420-8	Sequence 8, Appli
30	17.4	58.0	705	3	US-08-781-420-9	Sequence 9, Appli
31	17.4	58.0	705	3	US-08-874-102-8	Sequence 8, Appli
32	17.4	58.0	705	3	US-08-874-102-9	Sequence 9, Appli
c 33	17.4	58.0	705	3	US-08-984-919A-8	Sequence 8, Appli
34	17.4	58.0	705	3	US-08-984-919A-9	Sequence 9, Appli
c 35	17.4	58.0	705	3	US-09-006-595A-8	Sequence 8, Appli
36	17.4	58.0	705	3	US-09-006-595A-9	Sequence 9, Appli
c 37	17.4	58.0	707	3	US-08-781-420-5	Sequence 5, Appli
38	17.4	58.0	707	3	US-08-781-420-7	Sequence 7, Appli
c 39	17.4	58.0	707	3	US-08-874-102-5	Sequence 5, Appli
40	17.4	58.0	707	3	US-08-874-102-7	Sequence 7, Appli
c 41	17.4	58.0	707	3	US-08-984-919A-5	Sequence 5, Appli
42	17.4	58.0	707	3	US-08-984-919A-7	Sequence 7, Appli
c 43	17.4	58.0	707	3	US-09-006-595A-5	Sequence 5, Appli
44	17.4	58.0	707	3	US-09-006-595A-7	Sequence 7, Appli
45	17.4	58.0	993	3	US-09-489-039A-3737	Sequence 3737, Ap

ALIGNMENTS

```
RESULT 1
US-09-859-724-1
; Sequence 1, Application US/09859724
; Patent No. 6699843
; GENERAL INFORMATION:
;   APPLICANT: Pietras, Kristian
;   APPLICANT: Ostman, Arne
;   APPLICANT: Heldin, Carl-Henrik
;   APPLICANT: Rubin, Kristofer
;   TITLE OF INVENTION: METHOD FOR TREATMENT OF TUMORS USING NUCLEIC ACID
;   TITLE OF INVENTION: LIGANDS TO PDGF
;   FILE REFERENCE: NEX90
;   CURRENT APPLICATION NUMBER: US/09/859,724
;   CURRENT FILING DATE: 2001-09-21
;   PRIOR APPLICATION NUMBER: 60/205,006
;   PRIOR FILING DATE: 2000-05-17
;   PRIOR APPLICATION NUMBER: 08/479,725
;   PRIOR FILING DATE: 1995-06-07
;   PRIOR APPLICATION NUMBER: 08/479,783
;   PRIOR FILING DATE: 1995-06-07
;   PRIOR APPLICATION NUMBER: 08/618,693
;   PRIOR FILING DATE: 1996-03-20
;   PRIOR APPLICATION NUMBER: 08/991,743
;   PRIOR FILING DATE: 1997-12-16
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
;
; SEQ ID NO 1
; LENGTH: 30
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;   OTHER INFORMATION: Nucleic Acid Ligand
;   NAME/KEY: modified_base
;   LOCATION: (6)..(30)
;   OTHER INFORMATION: U at positions 6, 20 and 30 is
;   OTHER INFORMATION: 2'-fluoro-2'-deoxyuridine
;   NAME/KEY: modified_base
;   LOCATION: (8)..(29)
;   OTHER INFORMATION: C at positions 8, 21, 28, and 29 is
;   OTHER INFORMATION: 2'-fluoro-2'-deoxycytidine
;   NAME/KEY: modified_base
;   LOCATION: (9)..(31)
;   OTHER INFORMATION: G at positions 9, 15, 17, and 31 is
;   OTHER INFORMATION: 2'-O-Methyl-2'-deoxyguanosine
;   NAME/KEY: modified_base
;   LOCATION: (22)
;
```

```
;      OTHER INFORMATION: A at position 22 is 2'-O-Methyl-2'-deoxyadenosine
;      NAME/KEY: modified_base
;      LOCATION: (1)..(30)
;      OTHER INFORMATION: The residues at positions 9 and 10 are connected by a hexaethylen
;      OTHER INFORMATION: phosphoramidite linker. The residues at positions 21 and 22 are
;      OTHER INFORMATION: connected by a hexaethylene glycol phosphoramidite linker.
;      NAME/KEY: modified_base
;      LOCATION: (32)
;      OTHER INFORMATION: Nucleotide 32 is an inverted orientation T
;      OTHER INFORMATION: (3'-3'-linked)
US-09-859-724-1

  Query Match             100.0%;  Score 30;  DB 3;  Length 30;
  Best Local Similarity   100.0%;  Pred. No. 0.00017;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CAGGCUACGCGTAGAGCAUCAUGATCCUGT 30
              ||||||||||||||||||||||||||||||
Db          1 CAGGCUACGCGTAGAGCAUCAUGATCCUGT 30

RESULT 2
US-09-246-461-3
; Sequence 3, Application US/09246461
; Patent No. 6329145
; GENERAL INFORMATION:
;      APPLICANT: Janjic, Nebojsa
;      APPLICANT: Gold, Larry
;      TITLE OF INVENTION: Aptamers as Reagents for high Throughput Screening
;      FILE REFERENCE: NEX78
;      CURRENT APPLICATION NUMBER: US/09/246,461
;      CURRENT FILING DATE: 1999-02-09
;      NUMBER OF SEQ ID NOS: 4
;      SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;      LENGTH: 31
;      TYPE: DNA
;      ORGANISM: Artificial Sequence
;      FEATURE:
;      OTHER INFORMATION: Description of Artificial Sequence: Nucleic acid
;      OTHER INFORMATION: ligand
;      FEATURE:
;      NAME/KEY: modified_base
;      LOCATION: (1)..(31)
;      OTHER INFORMATION: The t at position 1 is 2'-NH2.  The u's and c's at
;      OTHER INFORMATION: positions 7, 9, 20, 21, 27, 28 and 29 are 2'-F.
;      OTHER INFORMATION: The a's and g's at positions 10, 15, 17, 22 and 30
;      OTHER INFORMATION: are 2'-O-methyl.
;      FEATURE:
;      NAME/KEY: modified_base
;      LOCATION: (1)..(31)
;      OTHER INFORMATION: The g at position 10 and the a at position 22 have
;      OTHER INFORMATION: been modified with polyethylene glycol of 18
;      OTHER INFORMATION: units.
US-09-246-461-3

  Query Match             100.0%;  Score 30;  DB 3;  Length 31;
  Best Local Similarity   100.0%;  Pred. No. 0.00017;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CAGGCUACGCGTAGAGCAUCAUGATCCUGT 30
              ||||||||||||||||||||||||||||||
Db          2 CAGGCUACGCGTAGAGCAUCAUGATCCUGT 31

RESULT 3
US-10-021-330-3
; Sequence 3, Application US/10021330
; Patent No. 6670132
; GENERAL INFORMATION:
;      APPLICANT: Janjic, Nebojsa
;      APPLICANT: Gold, Larry
;      TITLE OF INVENTION: Aptamers as Reagents for high Throughput Screening
;      FILE REFERENCE: NEX78
;      CURRENT APPLICATION NUMBER: US/10/021,330
;      CURRENT FILING DATE: 2001-12-10
;      PRIOR APPLICATION NUMBER: US/09/246,461
;      PRIOR FILING DATE: 1999-02-09
;      NUMBER OF SEQ ID NOS: 4
;      SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
;      LENGTH: 31
;      TYPE: DNA
;      ORGANISM: Artificial Sequence
;      FEATURE:
;      OTHER INFORMATION: Description of Artificial Sequence: Nucleic acid
;      OTHER INFORMATION: ligand
;      FEATURE:
;      NAME/KEY: modified_base
;      LOCATION: (1)..(31)
;      OTHER INFORMATION: The t at position 1 is 2'-NH2.  The u's and c's at
;      OTHER INFORMATION: positions 7, 9, 20, 21, 27, 28 and 29 are 2'-F.
;      OTHER INFORMATION: The a's and g's at positions 10, 15, 17, 22 and 30
;      OTHER INFORMATION: are 2'-O-methyl.
;      FEATURE:
;      NAME/KEY: modified_base
;      LOCATION: (1)..(31)
;      OTHER INFORMATION: The g at position 10 and the a at position 22 have
;      OTHER INFORMATION: been modified with polyethylene glycol of 18
;      OTHER INFORMATION: units.
US-10-021-330-3

  Query Match             100.0%;  Score 30;  DB 3;  Length 31;
  Best Local Similarity   100.0%;  Pred. No. 0.00017;
  Matches   30;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CAGGCUACGCGTAGAGCAUCAUGATCCUGT 30
              ||||||||||||||||||||||||||||||
Db          2 CAGGCUACGCGTAGAGCAUCAUGATCCUGT 31

RESULT 4
US-09-270-767-11347
; Sequence 11347, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
;      APPLICANT: Homburger et al.
;      TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
;      FILE REFERENCE: File Reference: 7326-094
;      CURRENT APPLICATION NUMBER: US/09/270,767
;      CURRENT FILING DATE: 1999-03-17
;      NUMBER OF SEQ ID NOS: 62517
;      SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11347
;      LENGTH: 606
;      TYPE: DNA
;      ORGANISM: Drosophila melanogaster
US-09-270-767-11347

  Query Match              65.3%;  Score 19.6;  DB 3;  Length 606;
  Best Local Similarity    84.6%;  Pred. No. 20;
  Matches   22;  Conservative    0;  Mismatches    4;  Indels    0;  Gaps    0;

Qy          1 CAGGCUACGCGTAGAGCAUCAUGATC 26
              ||||||||| || |||||||||||||
Db        279 CAGGCGACGGGTGGAGCACCATGATC 304

RESULT 5
US-09-270-767-14489
; Sequence 14489, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
;      APPLICANT: Homburger et al.
;      TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
;      FILE REFERENCE: File Reference: 7326-094
```

; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 14489
; LENGTH: 615
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
; FEATURE:
; OTHER INFORMATION: n means any nucleotide
US-09-270-767-14489

Query Match   65.3%; Score 19.6; DB 3; Length 615;
Best Local Similarity 84.6%; Pred. No. 21;
Matches   22; Conservative   0; Mismatches   0; Indels   4; Gaps   0;

Qy   1 CAGGCUACGCGTAGAGCAUCAUGATC 26
          |||||| ||||||||||| |||||
Db   117 CAGGCGACGCGGTGGAGCACCATGATC 142

RESULT 6
US-09-270-767-10626
; Sequence 10626, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
; APPLICANT: Homburger et al.
; TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
; FILE REFERENCE: File Reference: 7326-094
; CURRENT APPLICATION NUMBER: US/09/270,767
; CURRENT FILING DATE: 1999-03-17
; NUMBER OF SEQ ID NOS: 62517
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 10626
; LENGTH: 705
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
US-09-270-767-10626

Query Match   65.3%; Score 19.6; DB 3; Length 705;
Best Local Similarity 84.6%; Pred. No. 21;
Matches   22; Conservative   0; Mismatches   0; Indels   4; Gaps   0;

Qy   1 CAGGCUACGCGTAGAGCAUCAUGATC 26
          |||||| ||||||||||| |||||
Db   543 CAGGCGACGCGGTGGAGCACCATGATC 568

RESULT 7
US-09-949-016-16423
; Sequence 16423, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: CL001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16423
; LENGTH: 167708
; TYPE: DNA
; ORGANISM: Human
US-09-949-016-16423

Query Match   65.3%; Score 19.6; DB 3; Length 167708;
Best Local Similarity 73.1%; Pred. No. 65;
Matches   19; Conservative   3; Mismatches   4; Indels   0; Gaps   0;

Qy   5 CUACGCGTAGAGCAUCAUGATCCUGT 30
          |:||:||| ||||||:||:||| |:|:|
Db 50558 CTACCCGAAGAGCATCATGGTTCTGT 50583

RESULT 8
US-09-489-039A-2715
; Sequence 2715, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,039A
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,747
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 2715
; LENGTH: 603
; TYPE: DNA
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-2715

Query Match   63.3%; Score 19; DB 3; Length 603;
Best Local Similarity 74.1%; Pred. No. 39;
Matches   20; Conservative   2; Mismatches   2; Indels   5; Gaps   0;

Qy   3 GGCUACGCGTAGAGCAUCAUGATCCUG 29
          |||| ||| ||||||:|:|||||||||:|
Db   106 GGCCGACAGCTGGAGCATCATGATCCTG 132

RESULT 9
US-09-447-497-4
; Sequence 4, Application US/09447497
; Patent No. 6773911
; GENERAL INFORMATION:
; APPLICANT: Penninger, Josef M.
; APPLICANT: Kroemer, Guido P.
; APPLICANT: Siderovsi, David P.
; APPLICANT: Zamzami, Naoufal
; APPLICANT: Susin, Santos A.
; APPLICANT: Snow, Bryan E.L.
; TITLE OF INVENTION: APOPTOSIS-INDUCING FACTOR
; FILE REFERENCE: 01017/36780
; CURRENT APPLICATION NUMBER: US/09/447,497
; CURRENT FILING DATE: 1999-11-23
; EARLIER APPLICATION NUMBER: 60/109,595
; EARLIER FILING DATE: 1998-11-23
; NUMBER OF SEQ ID NOS: 25
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 1920
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(1875)
; FEATURE:
; OTHER INFORMATION: mouse apoptosis-inducing factor (mAIF isoform);
; OTHER INFORMATION: mAIF-alt-exon-Gold
US-09-447-497-4

Query Match   62.7%; Score 18.8; DB 3; Length 1920;
Best Local Similarity 81.8%; Pred. No. 61;
Matches   18; Conservative   2; Mismatches   2; Indels   0; Gaps   0;

```
Qy        8 CGCGTAGAGCAUCATGATCCUG 29
            || ||||||||||:||||||| :|
Db     1387 CGAGTAGAGCATCATGATCATG 1408

RESULT 10
US-09-447-497-1
; Sequence 1, Application US/09447497
; Patent No. 6773911
; GENERAL INFORMATION:
;   APPLICANT: Penninger, Josef M.
;   APPLICANT: Kroemer, Guido P.
;   APPLICANT: Siderovsi, David P.
;   APPLICANT: Zamzami, Naoufal
;   APPLICANT: Susin, Santos A.
;   APPLICANT: Snow, Bryan E.L.
;   TITLE OF INVENTION: APOPTOSIS-INDUCING FACTOR
;   FILE REFERENCE: 01017/36780
;   CURRENT APPLICATION NUMBER: US/09/447,497
;   CURRENT FILING DATE: 1999-11-23
;   EARLIER APPLICATION NUMBER: 60/109,595
;   EARLIER FILING DATE: 1998-11-23
;   NUMBER OF SEQ ID NOS: 25
;   SOFTWARE: PatentIn Ver. 2.0
;
; SEQ ID NO 1
;   LENGTH: 1932
;   TYPE: DNA
;   ORGANISM: Mus musculus
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: (1)..(1887)
;   FEATURE:
;   OTHER INFORMATION: mouse apoptosis-inducing factor (mAIF) ;
;   OTHER INFORMATION: mAIF-Gold.B1B
US-09-447-497-1

Query Match        62.7%;    Score 18.8;  DB 3;  Length 1932;
Best Local Similarity 81.8%;  Pred. No. 61;
Matches   18; Conservative   2; Mismatches   2; Indels   2; Gaps   0;

Qy        8 CGCGTAGAGCAUCATGATCCUG 29
            || ||||||||||:||||||| :|
Db     1399 CGAGTAGAGCATCATGATCATG 1420

RESULT 11
US-09-533-559-2284
; Sequence 2284, Application US/09533559
; Patent No. 6902887
; GENERAL INFORMATION:
;   APPLICANT: Randy M. Berka
;   APPLICANT: Michael W. Rey
;   APPLICANT: Jeffrey R.Shuster
;   APPLICANT: Sakari Kauppinen
;   APPLICANT: Ib Groth Clausen
;   APPLICANT: Peter Bjarke Olsen
;   TITLE OF INVENTION: Methods For Monitoring Multiple Gene
;   TITLE OF INVENTION: Expression
;   FILE REFERENCE: 5849.200-US
;   CURRENT APPLICATION NUMBER: US/09/533,559
;   CURRENT FILING DATE: 2000-03-22
;   EARLIER APPLICATION NUMBER: 09/273,623
;   EARLIER FILING DATE: 1999-03-22
;   NUMBER OF SEQ ID NOS: 7860
;   SOFTWARE: FastSEQ for Windows Version 4.0
;
; SEQ ID NO 2284
;   LENGTH: 485
;   TYPE: DNA
;   ORGANISM: Fusarium venenatum
;   FEATURE:
;   NAME/KEY: misc_feature
;   LOCATION: (1)...(485)
;   OTHER INFORMATION: n = A,T,C or G
```

```
US-09-533-559-2284

Query Match        61.3%;    Score 18.4;  DB 3;  Length 485;
Best Local Similarity 71.4%;  Pred. No. 70;
Matches   20; Conservative   2; Mismatches   6; Indels   0; Gaps   0;

Qy        2 AGGCUACGGCGTAGAGCAUCATGATCCUG 29
            |||| || |:|||||||||:||||||||:|
Db      181 AGGCAGCGTTTGGGGCATCATGATCCTG 208

RESULT 12
US-08-753-247-7/c
; Sequence 7, Application US/08753247
; Patent No. 6210929
; GENERAL INFORMATION:
;   APPLICANT: SCHLOKAT, Uwe
;   APPLICANT: FISCHER, Bernhard
;   APPLICANT: FALKNER, Falko-Guenther
;   APPLICANT: DORNER, Friedrich
;   APPLICANT: EIBL, Johann
;   TITLE OF INVENTION:  A FUSION PROTEIN COMPRISING A FURIN
;   TITLE OF INVENTION:  DERIVATIVE OR A DERIVATIVE OF A FURIN ANALOGUE AND A
;   TITLE OF INVENTION:  HETEROLOGOUS SEQUENCE
;   NUMBER OF SEQUENCES:   29
;   CORRESPONDENCE ADDRESS:
;     ADDRESSEE:  Foley & Lardner
;     STREET: 3000 K Street, N.W., Suite 500
;     CITY:  Washington
;     STATE:  D.C.
;     COUNTRY: USA
;     ZIP: 20007-5109
;   COMPUTER READABLE FORM:
;     MEDIUM TYPE: Floppy disk
;     COMPUTER:  IBM PC compatible
;     OPERATING SYSTEM:  PC-DOS/MS-DOS
;     SOFTWARE:  PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;     APPLICATION NUMBER: US/08/753,247
;     FILING DATE:  22-NOV-1996
;     CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;     APPLICATION NUMBER:  AT 1928/95
;     FILING DATE:  24-NOV-1995
;   ATTORNEY/AGENT INFORMATION:
;     NAME:  BENT, Stephen A.
;     REGISTRATION NUMBER: 29,768
;     REFERENCE/DOCKET NUMBER:  40433/149
;   TELECOMMUNICATION INFORMATION:
;     TELEPHONE:  (202)672-5300
;     TELEFAX:  (202)672-5399
;     TELEX:  904136
; INFORMATION FOR SEQ ID NO:  7:
;   SEQUENCE CHARACTERISTICS:
;     LENGTH:  50 base pairs
;     TYPE:  nucleic acid
;     STRANDEDNESS:  single
;     TOPOLOGY:  linear
;   MOLECULE TYPE: DNA (genomic)
US-08-753-247-7

Query Match        60.0%;    Score 18;    DB 3;  Length 50;
Best Local Similarity 69.2%;  Pred. No. 68;
Matches   18; Conservative   3; Mismatches   5; Indels   0; Gaps   0;

Qy        3 GGCUACGGCGTAGAGCAUCATGATCCU 28
            ||| :|||| ||||||||| ||| :|||
Db       42 GGCTGCGCGCCAGGGCATCATCATCAT 17

RESULT 13
US-09-270-767-8822
; Sequence 8822, Application US/09270767
```

Patent No. 6703491
; GENERAL INFORMATION:
;   APPLICANT: Homburger et al.
;   TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
;   FILE REFERENCE: File Reference: 7326-094
;   CURRENT APPLICATION NUMBER: US/09/270,767
;   CURRENT FILING DATE: 1999-03-17
;   NUMBER OF SEQ ID NOS: 62517
;   SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 8822
;   LENGTH: 477
;   TYPE: DNA
;   ORGANISM: Drosophila melanogaster
US-09-270-767-8822

Query Match   60.0%; Score 18; DB 3; Length 477;
Best Local Similarity 73.1%; Pred. No. 1.1e+02;
Matches   19; Conservative   2; Mismatches   5; Indels   0; Gaps   0;

Qy   2 AGGCUACGCGUAGAGCAUCAUGAUCC   27
         ||||:| ||||| ||||| || |||
Db 339 AGGCTATGCGTATAGCATGATTAACC 364

RESULT 14
US-09-270-767-24104
; Sequence 24104, Application US/09270767
; Patent No. 6703491
; GENERAL INFORMATION:
;   APPLICANT: Homburger et al.
;   TITLE OF INVENTION: Nucleic acids and proteins of Drosophila melanogaster
;   FILE REFERENCE: File Reference: 7326-094
;   CURRENT APPLICATION NUMBER: US/09/270,767
;   CURRENT FILING DATE: 1999-03-17
;   NUMBER OF SEQ ID NOS: 62517
;   SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 24104
;   LENGTH: 477
;   TYPE: DNA
;   ORGANISM: Drosophila melanogaster
US-09-270-767-24104

Query Match   60.0%; Score 18; DB 3; Length 477;
Best Local Similarity 73.1%; Pred. No. 1.1e+02;
Matches   19; Conservative   2; Mismatches   5; Indels   0; Gaps   0;

Qy   2 AGGCUACGCGUAGAGCAUCAUGAUCC   27
         ||||:| ||||| ||||| || |||
Db 339 AGGCTATGCGTATAGCATGATTAACC 364

RESULT 15
US-08-753-247-8
; Sequence 8, Application US/08753247
; Patent No. 6210929
; GENERAL INFORMATION:
;   APPLICANT: SCHLOKAT, Uwe
;   APPLICANT: FISCHER, Bernhard
;   APPLICANT: FALKNER, Falko-Guenther
;   APPLICANT: DORNER, Friedrich
;   APPLICANT: EIBL, Johann
;   TITLE OF INVENTION: A FUSION PROTEIN COMPRISING A FURIN
;   TITLE OF INVENTION: DERIVATIVE OR A DERIVATIVE OF A FURIN ANALOGUE AND A
;   TITLE OF INVENTION: HETEROLOGOUS SEQUENCE
;   NUMBER OF SEQUENCES: 29
;   CORRESPONDENCE ADDRESS:
;   ADDRESSEE: Foley & Lardner
;   STREET: 3000 K Street, N.W., Suite 500
;   CITY: Washington
;   STATE: D.C.
;   COUNTRY: USA
;   ZIP: 20007-5109
;   COMPUTER READABLE FORM:

;   MEDIUM TYPE: Floppy disk
;   COMPUTER: IBM PC compatible
;   OPERATING SYSTEM: PC-DOS/MS-DOS
;   SOFTWARE: PatentIn Release #1.0, Version #1.30
;   CURRENT APPLICATION DATA:
;   APPLICATION NUMBER: US/08/753,247
;   FILING DATE: 22-NOV-1996
;   CLASSIFICATION: 435
;   PRIOR APPLICATION DATA:
;   APPLICATION NUMBER: AT 1928/95
;   FILING DATE: 24-NOV-1995
;   ATTORNEY/AGENT INFORMATION:
;   NAME: BENT, Stephen A.
;   REGISTRATION NUMBER: 29,768
;   REFERENCE/DOCKET NUMBER: 40433/149
;   TELECOMMUNICATION INFORMATION:
;   TELEPHONE: (202)672-5300
;   TELEFAX: (202)672-5399
;   TELEX: 904136
; INFORMATION FOR SEQ ID NO: 8:
;   SEQUENCE CHARACTERISTICS:
;   LENGTH: 2142 base pairs
;   TYPE: nucleic acid
;   STRANDEDNESS: single
;   TOPOLOGY: linear
;   MOLECULE TYPE: DNA (genomic)
;   FEATURE:
;   NAME/KEY: CDS
;   LOCATION: 1..2139
;   FEATURE:
;   NAME/KEY: mat_peptide
;   LOCATION: 1..2139
US-08-753-247-8

Query Match   60.0%; Score 18; DB 3; Length 2142;
Best Local Similarity 69.2%; Pred. No. 1.5e+02;
Matches   18; Conservative   3; Mismatches   5; Indels   0; Gaps   0;

Qy   3 GGCUACGCGUAGAGCAUCAUGAUCCU   28
          |||:| |||||| || ||||| ||| |||
Db 2108 GGCTGCGCGCCAGGGCATCATCATCAT 2133

Search completed: January 8, 2006, 17:08:34
Job time : 103 secs

```
              GenCore version 5.1.6
         Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:    January 8, 2006, 15:50:54 ; Search time 560 Seconds
                                       (without alignments)
                                       443.002 Million cell updates/sec

Title:          US-10-791-367-1
Perfect score:  30
Sequence:       1 caggcuacgcgtagagcaucatgatccugt 30

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       979542 seqs, 413489005 residues

Total number of hits satisfying chosen parameters:    19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :   Published Applications NA Main:*
           1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
           2: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
           3: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
           4: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
           5: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
           6: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
           7: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
           8: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
           9: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
          10: /cgn2_6/ptodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.
```

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	30	100.0	30	3	US-09-859-724-1	Sequence 1, Appli
2	30	100.0	30	8	US-10-791-367-1	Sequence 1, Appli
3	30	100.0	30	9	US-10-926-806-21	Sequence 21, Appl
4	30	100.0	31	5	US-10-021-330-3	Sequence 3, Appli
5	30	100.0	31	8	US-10-746-565-3	Sequence 3, Appli
6	29	96.7	29	9	US-10-829-504A-8	Sequence 8, Appli
7	29	96.7	29	9	US-10-873-853-8	Sequence 8, Appli
8	29	96.7	29	9	US-10-980-211-8	Sequence 8, Appli
9	19.8	66.0	2862		US-10-282-122A-41310	Sequence 41310, A
c 10	19.6	65.3	2349	10	US-11-097-143-16217	Sequence 16217, A
c 11	19.6	65.3	4992	10	US-11-097-143-16216	Sequence 16216, A
12	19.4	64.7	403	8	US-10-425-115-17734	Sequence 17734, A
13	19.4	64.7	1335	8	US-10-425-115-139288	Sequence 139288,
14	19.2	64.0	466	7	US-10-767-701-29066	Sequence 29066, A
c 15	19	63.3	663	8	US-10-762-107-27	Sequence 27, Appl
c 16	19	63.3	734	7	US-10-425-114-34193	Sequence 34193, A
c 17	19	63.3	734	8	US-10-425-115-145955	Sequence 145955,
18	19	63.3	1359	7	US-10-437-963-24608	Sequence 24608, A
c 19	19	63.3	36602	8	US-10-762-107-1	Sequence 1, Appli
20	19	63.3	2256646	7	US-10-470-565-1	Sequence 1, Appli
21	18.8	62.7	396	7	US-10-424-599-92411	Sequence 92411, A
22	18.8	62.7	1617	8	US-10-481-113-4	Sequence 4, Appli
23	18.8	62.7	1638	8	US-10-481-113-5	Sequence 5, Appli
24	18.8	62.7	1857	7	US-10-437-963-74656	Sequence 74656, A
25	18.8	62.7	1920	6	US-10-308-936-4	Sequence 4, Appli
26	18.8	62.7	1932	8	US-10-308-936-1	Sequence 1, Appli
27	18.8	62.7	1932	7	US-10-419-256-1	Sequence 1, Appli
c 28	18.8	62.7	6899	3	US-09-938-842A-2030	Sequence 2030, Ap
c 29	18.8	62.7	6899	3	US-09-938-842A-2030	Sequence 2030, Ap
c 30	18.8	62.7	40633	5	US-10-087-192-1	Sequence 1, Appli
31	18.6	62.0	573	7	US-10-021-323-3107	Sequence 3107, Ap
c 32	18.6	62.0	652	7	US-10-021-323-1390	Sequence 1390, Ap
c 33	18.6	62.0	1430	8	US-10-767-795-86	Sequence 86, Appl
34	18.6	62.0	1944	7	US-10-437-963-31340	Sequence 31340, A
c 35	18.6	62.0	12195	9	US-10-661-398-13	Sequence 13, Appl
c 36	18.6	62.0	16061	9	US-10-661-398-12	Sequence 12, Appl
c 37	18.4	61.3	147	8	US-10-425-115-150217	Sequence 150217,
38	18.4	61.3	353	9	US-09-922-293-2640	Sequence 2640, Ap
39	18.4	61.3	485	8	US-10-653-047-2284	Sequence 2284, Ap
c 40	18.4	61.3	602	4	US-09-925-065A-422787	Sequence 422787,
41	18.4	61.3	646	6	US-10-027-632-319063	Sequence 319063,
42	18.4	61.3	646	6	US-10-027-632-319063	Sequence 319063,
43	18	60.0	541	4	US-09-925-065A-662266	Sequence 662266,
c 44	18	60.0	602	4	US-09-925-065A-788924	Sequence 788924,
c 45	18	60.0	602	4	US-09-925-065A-847051	Sequence 847051,

ALIGNMENTS

```
RESULT 1
US-09-859-724-1
; Sequence 1, Application US/09859724
; Patent No. US20020034506A1
; GENERAL INFORMATION:
;   APPLICANT: Pietras, Kristian
;   APPLICANT: Ostman, Arne
;   APPLICANT: Heldin, Carl-Henrik
;   APPLICANT: Rubin, Kristofer
;   TITLE OF INVENTION: METHOD FOR TREATMENT OF TUMORS USING NUCLEIC ACID
;   TITLE OF INVENTION: LIGANDS TO PDGF
;   FILE REFERENCE: NEX90
;   CURRENT APPLICATION NUMBER: US/09/859,724
;   CURRENT FILING DATE: 2001-09-21
;   PRIOR APPLICATION NUMBER: 60/205,006
;   PRIOR FILING DATE: 2000-05-17
;   PRIOR APPLICATION NUMBER: 08/479,725
;   PRIOR FILING DATE: 1995-06-07
;   PRIOR APPLICATION NUMBER: 08/479,783
;   PRIOR FILING DATE: 1995-06-07
;   PRIOR APPLICATION NUMBER: 08/618,693
;   PRIOR FILING DATE: 1996-03-20
;   PRIOR APPLICATION NUMBER: 08/991,743
;   PRIOR FILING DATE: 1997-12-16
;   NUMBER OF SEQ ID NOS: 2
;   SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 30
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;     OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;     OTHER INFORMATION: Nucleic Acid Ligand
;     NAME/KEY: modified_base
;     LOCATION: (6)..(30)
;     OTHER INFORMATION: U at positions 6, 20 and 30 is
;     OTHER INFORMATION: 2'-fluoro-2'-deoxyuridine
;     NAME/KEY: modified_base
;     LOCATION: (8)..(29)
;     OTHER INFORMATION: C at positions 8, 21, 28, and 29 is
;     OTHER INFORMATION: 2'-fluoro-2'-deoxycytidine
;     NAME/KEY: modified_base
;     LOCATION: (9)..(31)
;     OTHER INFORMATION: G at positions 9, 15, 17, and 31 is
;     OTHER INFORMATION: 2'-0-Methyl-2'-deoxyguanosine
;     NAME/KEY: modified_base
```

```
;      LOCATION: (22)
;      OTHER INFORMATION: A at position 22 is 2'-O-Methyl-2'-deoxyadenosine
;      NAME/KEY: modified_base
;      LOCATION: (1)..(30)
;      OTHER INFORMATION: The residues at positions 9 and 10 are connected by a hexaethylen
;      OTHER INFORMATION: phosphoramidite linker. The residues at positions 21 and 22 are
;      NAME/KEY: modified_base
;      OTHER INFORMATION: connected by a hexaethylene glycol phosphoramidite linker.
;      LOCATION: (32)
;      OTHER INFORMATION: Nucleotide 32 is an inverted orientation T
;      OTHER INFORMATION: (3'-3'-linked)
US-09-859-724-1

Query Match             100.0%;  Score 30;  DB 3;  Length 30;
Best Local Similarity   100.0%;  Pred. No. 0.00046;
Matches    30;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          1 CAGGCUACGCGTAGAGCAUCATGATCCUGT 30
              ||||||||||||||||||||||||||||||
Db          1 CAGGCUACGCGTAGAGCAUCATGATCCUGT 30

RESULT 2
US-10-791-367-1
; Sequence 1, Application US/10791367
; Publication No. US20040224335A1
; GENERAL INFORMATION:
;   APPLICANT: Pietras, Kristian
;   APPLICANT: Ostman, Arne
;   APPLICANT: Heldin, Carl-Henrik
;   APPLICANT: Rubin, Kristofer
;   TITLE OF INVENTION: METHOD FOR TREATMENT OF TUMORS USING NUCLEIC ACID
;   FILE REFERENCE: NEX90
;   CURRENT APPLICATION NUMBER: US/10/791,367
;   CURRENT FILING DATE: 2004-03-02
;   PRIOR APPLICATION NUMBER: US/09/859,724
;   PRIOR FILING DATE: 2001-09-21
;   PRIOR APPLICATION NUMBER: 60/205,006
;   PRIOR FILING DATE: 2000-05-17
;   PRIOR APPLICATION NUMBER: 08/479,725
;   PRIOR FILING DATE: 1995-06-07
;   PRIOR APPLICATION NUMBER: 08/479,783
;   PRIOR FILING DATE: 1995-06-07
;   PRIOR APPLICATION NUMBER: 08/618,693
;   PRIOR FILING DATE: 1996-03-20
;   PRIOR APPLICATION NUMBER: 08/991,743
;   PRIOR FILING DATE: 1997-12-16
;   NUMBER OF SEQ ID NOS: 2
;   SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
;   LENGTH: 30
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;      OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;      OTHER INFORMATION: Nucleic Acid Ligand
;   FEATURE:
;      NAME/KEY: modified_base
;      LOCATION: (6)..(30)
;      OTHER INFORMATION: U at positions 6, 20 and 30 is
;      OTHER INFORMATION: 2'-fluoro-2'-deoxyuridine
;   FEATURE:
;      NAME/KEY: modified_base
;      LOCATION: (8)..(29)
;      OTHER INFORMATION: C at positions 8, 21, 28, and 29 is
;      OTHER INFORMATION: 2'-fluoro-2'-deoxycytidine
;   FEATURE:
;      NAME/KEY: modified_base
;      LOCATION: (9)..(31)
;      OTHER INFORMATION: G at positions 9, 15, 17, and 31 is
;      OTHER INFORMATION: 2'-O-Methyl-2'-deoxyguanosine
;   FEATURE:

;      NAME/KEY: modified_base
;      LOCATION: (22)
;      OTHER INFORMATION: A at position 22 is 2'-O-Methyl-2'-deoxyadenosine
;   FEATURE:
;      NAME/KEY: modified_base
;      LOCATION: (1)..(30)
;      OTHER INFORMATION: The residues at positions 9 and 10 are connected by a hexaethylene
;      OTHER INFORMATION: phosphoramidite linker. The residues at positions 21 and 22 are
;      OTHER INFORMATION: connected by a hexaethylene glycol phosphoramidite linker.
;   FEATURE:
;      NAME/KEY: modified_base
;      LOCATION: (32)
;      OTHER INFORMATION: Nucleotide 32 is an inverted orientation T
;      OTHER INFORMATION: (3'-3'-linked)
US-10-791-367-1

Query Match             100.0%;  Score 30;  DB 8;  Length 30;
Best Local Similarity   100.0%;  Pred. No. 0.00046;
Matches    30;  Conservative   0;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          1 CAGGCUACGCGTAGAGCAUCATGATCCUGT 30
              ||||||||||||||||||||||||||||||
Db          1 CAGGCUACGCGTAGAGCAUCATGATCCUGT 30

RESULT 3
US-10-926-806-21
; Sequence 21, Application US/10926806
; Publication No. US2005009257A1
; GENERAL INFORMATION:
;   APPLICANT: SHIMA, DAVID
;   APPLICANT: CALIAS, PERRY
;   APPLICANT: ADAMIS, ANTHONY P.
;   TITLE OF INVENTION: COMBINATION THERAPY FOR THE TREATMENT OF OCULAR
;   TITLE OF INVENTION: NEOVASCULAR DISORDERS
;   FILE REFERENCE: 112089.184 EYE-013
;   CURRENT APPLICATION NUMBER: US/10/926,806
;   CURRENT FILING DATE: 2004-08-26
;   PRIOR APPLICATION NUMBER: 60/556,837
;   PRIOR FILING DATE: 2004-03-26
;   PRIOR APPLICATION NUMBER: 60/498,407
;   PRIOR FILING DATE: 2003-08-27
;   NUMBER OF SEQ ID NOS: 22
;   SOFTWARE: PatentIn Ver. 3.2
; SEQ ID NO 21
;   LENGTH: 30
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;      OTHER INFORMATION: Description of Artificial Sequence: Synthetic
;      OTHER INFORMATION: anti-PDGF aptamer
US-10-926-806-21

Query Match             100.0%;  Score 30;  DB 9;  Length 30;
Best Local Similarity    90.0%;  Pred. No. 0.00046;
Matches    27;  Conservative   3;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          1 CAGGCUACGCGTAGAGCAUCATGATCCUGT 30
              |||||||||||||||||:||||||||||:||
Db          1 CAGGCUACGCGTAGAGCATCATGATCCTGT 30

RESULT 4
US-10-021-330-3
; Sequence 3, Application US/10021330
; Publication No. US2002019617A1
; GENERAL INFORMATION:
;   APPLICANT: Janjic, Nebojsa
;   APPLICANT: Gold, Larry
;   TITLE OF INVENTION: Aptamers as Reagents for high Throughput Screening
;   FILE REFERENCE: NEX78
;   CURRENT APPLICATION NUMBER: US/10/021,330
;   CURRENT FILING DATE: 2001-12-10
```

```
; PRIOR APPLICATION NUMBER: US/09/246,461
; PRIOR FILING DATE: 1999-02-09
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 31
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Nucleic acid
; OTHER INFORMATION: ligand
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1)..(31)
; OTHER INFORMATION: The t at position 1 is 2'-NH2.  The u's and c's at
; OTHER INFORMATION: positions 7, 9, 20, 21, 27, 28 and 29 are 2'-F.
; OTHER INFORMATION: The a's and g's at positions 10, 15, 17, 22 and 30
; OTHER INFORMATION: are 2'-O-methyl.
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1)..(31)
; OTHER INFORMATION: The g at position 10 and the a at position 22 have
; OTHER INFORMATION: been modified with polyethylene glycol of 18
; OTHER INFORMATION: units.
US-10-021-330-3


Query Match          100.0%;   Score 30;  DB 5;  Length 31;
Best Local Similarity 100.0%;  Pred. No. 0.00046;
Matches  30;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 CAGGCUACGCGTAGAGCAUCATGATCCUGT 30
        ||||||||||||||||||||||||||||||
Db    2 CAGGCUACGCGTAGAGCAUCATGATCCUGT 31


RESULT 5
US-10-746-565-3
; Sequence 3, Application US/10746565
; Publication No. US20050048521A1
; GENERAL INFORMATION:
; APPLICANT: Janjic, Nebojsa
; APPLICANT: Gold, Larry
; TITLE OF INVENTION: Aptamers as Reagents for high Throughput Screening
; FILE REFERENCE: NEX78
; CURRENT APPLICATION NUMBER: US/10/746,565
; CURRENT FILING DATE:   2003-12-24
; PRIOR APPLICATION NUMBER: US/10/021,330
; PRIOR FILING DATE: 2001-12-10
; PRIOR APPLICATION NUMBER: US/09/246,461
; PRIOR FILING DATE: 1999-02-09
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3
; LENGTH: 31
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Nucleic acid
; OTHER INFORMATION: ligand
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1)..(31)
; OTHER INFORMATION: The t at position 1 is 2'-NH2.  The u's and c's at
; OTHER INFORMATION: positions 7, 9, 20, 21, 27, 28 and 29 are 2'-F.
; OTHER INFORMATION: The a's and g's at positions 10, 15, 17, 22 and 30
; OTHER INFORMATION: are 2'-O-methyl.
; FEATURE:
; NAME/KEY: modified_base
; LOCATION: (1)..(31)
; OTHER INFORMATION: The g at position 10 and the a at position 22 have
; OTHER INFORMATION: been modified with polyethylene glycol of 18
; OTHER INFORMATION: units.
US-10-746-565-3


Query Match          100.0%;   Score 30;  DB 8;  Length 31;
Best Local Similarity 100.0%;  Pred. No. 0.00046;
Matches  30;  Conservative  0;  Mismatches  0;  Indels  0;  Gaps  0;

Qy    1 CAGGCUACGCGTAGAGCAUCATGATCCUGT 30
        ||||||||||||||||||||||||||||||
Db    2 CAGGCUACGCGTAGAGCAUCATGATCCUGT 31


RESULT 6
US-10-829-504A-8
; Sequence 8, Application US/10829504A
; Publication No. US20040253679A1
; GENERAL INFORMATION:
; APPLICANT: Epstein, David
; APPLICANT: Grate, Dilara
; APPLICANT: Stanton, Martin
; APPLICANT: Diener, John L.
; APPLICANT: Wilson, Charles
; APPLICANT: McCauley, Thomas
; APPLICANT: DeSouza, Errol
; TITLE OF INVENTION: Stabilized Aptamers to Platelet Derived Growth Factor and Their
; TITLE OF INVENTION: Use as Oncology Therapeutics
; FILE REFERENCE: 23239-558A
; CURRENT APPLICATION NUMBER: US/10/829,504A
; CURRENT FILING DATE: 2004-04-21
; PRIOR APPLICATION NUMBER: 10/762915
; PRIOR FILING DATE: 2004-01-21
; PRIOR APPLICATION NUMBER: 60/441357
; PRIOR FILING DATE: 2003-01-21
; PRIOR APPLICATION NUMBER: 60/463095
; PRIOR FILING DATE: 2003-04-15
; PRIOR APPLICATION NUMBER: 60/464179
; PRIOR FILING DATE: 2003-04-21
; PRIOR APPLICATION NUMBER: 60/465055
; PRIOR FILING DATE: 2003-04-23
; PRIOR APPLICATION NUMBER: 60/469628
; PRIOR FILING DATE: 2003-05-08
; PRIOR APPLICATION NUMBER: 60/474680
; PRIOR FILING DATE: 2003-05-29
; PRIOR APPLICATION NUMBER: 60/491019
; PRIOR FILING DATE: 2003-07-29
; PRIOR APPLICATION NUMBER: 60/512071
; PRIOR FILING DATE: 2003-10-17
; PRIOR APPLICATION NUMBER: 60/537201
; PRIOR FILING DATE: 2004-01-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 98
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 8
; LENGTH: 29
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: aptamer
US-10-829-504A-8


Query Match          96.7%;   Score 29;  DB 8;  Length 29;
Best Local Similarity 89.7%;  Pred. No. 0.0014;
Matches  26;  Conservative  3;  Mismatches  3;  Indels  0;  Gaps  0;

Qy    1 CAGGCUACGCGTAGAGCAUCATGATCCUG 29
        |||||:||||||||||||||:|||||::|
Db    1 CAGGCTACGCGTAGAGCATCATGATCCTG 29


RESULT 7
US-10-873-853-8
; Sequence 8, Application US/10873853
; Publication No. US20050124565A1
; GENERAL INFORMATION:
; APPLICANT: Epstein, David
```

APPLICANT: Grate, Dilara
APPLICANT: Stanton, Martin
APPLICANT: Diener, John
APPLICANT: Wilson, Charles
APPLICANT: McCauley, Thomas
APPLICANT: DeSouza, Errol
TITLE OF INVENTION: Stabilized Aptamers to Platelet Derived Growth Factor and Their
TITLE OF INVENTION: Use as Oncology Therapeutics
FILE REFERENCE: 23339-558A CIP
CURRENT APPLICATION NUMBER: US/10/873,853
CURRENT FILING DATE: 2004-06-21
PRIOR APPLICATION NUMBER: 10/829504
PRIOR FILING DATE: 2004-04-21
PRIOR APPLICATION NUMBER: 10/762915
PRIOR FILING DATE: 2004-01-21
PRIOR APPLICATION NUMBER: 10/718833
PRIOR FILING DATE: 2003-11-21
PRIOR APPLICATION NUMBER: 60/441357
PRIOR FILING DATE: 2003-01-21
PRIOR APPLICATION NUMBER: 60/463095
PRIOR FILING DATE: 2003-04-15
PRIOR APPLICATION NUMBER: 60/428102
PRIOR FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: 60/464179
PRIOR FILING DATE: 2003-04-21
PRIOR APPLICATION NUMBER: 60/465055
PRIOR FILING DATE: 2003-04-23
PRIOR APPLICATION NUMBER: 60/512071
PRIOR FILING DATE: 2003-10-17
PRIOR APPLICATION NUMBER: 60/537201
PRIOR FILING DATE: 2004-01-16
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 98
SOFTWARE: PatentIn version 3.2
SEQ ID NO 8
  LENGTH: 29
  TYPE: DNA
  ORGANISM: Artificial
  FEATURE:
  OTHER INFORMATION: aptamer
US-10-873-853-8

Query Match   96.7%;   Score 29; DB 9; Length 29;
Best Local Similarity 89.7%; Pred. No. 0.0014;
Matches   26; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```
Qy  1  CAGGCUACGCGTAGAGCAUCATGATCCUG 29
       |||||:|||||||||||||:||||||::|
Db  1  CAGGCTACGCGTAGAGCATCATGATCCTG 29
```

RESULT 8
US-10-980-211-8
; Sequence 8, Application US/10980211
; Publication No. US2005015935A1
; GENERAL INFORMATION:
APPLICANT: Grate, Dilara
APPLICANT: Diener, John
APPLICANT: Wilson, Charles
APPLICANT: McCauley, Thomas
TITLE OF INVENTION: Stabilized Aptamers to Platelet Derived Growth Factor and Their
TITLE OF INVENTION: Use as Oncology Therapeutics
FILE REFERENCE: 23339-558A CIP2
CURRENT APPLICATION NUMBER: US/10/980,211
CURRENT FILING DATE: 2004-11-02
PRIOR APPLICATION NUMBER: 10/873853
PRIOR FILING DATE: 2004-06-21
PRIOR APPLICATION NUMBER: 10/829504
PRIOR FILING DATE: 2004-04-21
PRIOR APPLICATION NUMBER: 10/762915
PRIOR FILING DATE: 2004-01-21
PRIOR APPLICATION NUMBER: 10/718833
PRIOR FILING DATE: 2003-11-21

PRIOR APPLICATION NUMBER: 60/441357
PRIOR FILING DATE: 2003-01-21
PRIOR APPLICATION NUMBER: 60/463095
PRIOR FILING DATE: 2003-04-15
PRIOR APPLICATION NUMBER: 60/428102
PRIOR FILING DATE: 2002-11-21
PRIOR APPLICATION NUMBER: 60/464179
PRIOR FILING DATE: 2003-04-21
PRIOR APPLICATION NUMBER: 60/465055
PRIOR FILING DATE: 2003-04-23
PRIOR APPLICATION NUMBER: 60/512071
PRIOR FILING DATE: 2003-10-17
Remaining Prior Application data removed - See File Wrapper or PALM.
NUMBER OF SEQ ID NOS: 98
SOFTWARE: PatentIn version 3.2
SEQ ID NO 8
  LENGTH: 29
  TYPE: DNA
  ORGANISM: Artificial
  FEATURE:
  OTHER INFORMATION: aptamer
US-10-980-211-8

Query Match   96.7%;   Score 29; DB 9; Length 29;
Best Local Similarity 89.7%; Pred. No. 0.0014;
Matches   26; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

```
Qy  1  CAGGCUACGCGTAGAGCAUCATGATCCUG 29
       |||||:|||||||||||||:||||||::|
Db  1  CAGGCTACGCGTAGAGCATCATGATCCTG 29
```

RESULT 9
US-10-282-122A-41310
; Sequence 41310, Application US/10282122A
; Publication No. US2004002912 9A1
; GENERAL INFORMATION:
APPLICANT: Wang, Liangsu
APPLICANT: Zamudio, Carlos
APPLICANT: Malone, Cheryl
APPLICANT: Haselbeck, Robert
APPLICANT: Ohlsen, Kari
APPLICANT: Zyskind, Judith
APPLICANT: Wall, Daniel
APPLICANT: Trawick, John
APPLICANT: Carr, Grant
APPLICANT: Yamamoto, Robert
APPLICANT: Forsyth, R.
APPLICANT: Xu, H.
TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
FILE REFERENCE: ELITRA.034A
CURRENT APPLICATION NUMBER: US/10/282,122A
CURRENT FILING DATE: 2003-02-20
PRIOR APPLICATION NUMBER: 60/191,078
PRIOR FILING DATE: 2000-03-21
PRIOR APPLICATION NUMBER: 60/206,848
PRIOR FILING DATE: 2000-05-23
PRIOR APPLICATION NUMBER: 60/207,727
PRIOR FILING DATE: 2000-05-26
PRIOR APPLICATION NUMBER: 60/230,335
PRIOR FILING DATE: 2000-09-06
PRIOR APPLICATION NUMBER: 60/230,347
PRIOR FILING DATE: 2000-09-09
PRIOR APPLICATION NUMBER: 60/242,578
PRIOR FILING DATE: 2000-10-23
PRIOR APPLICATION NUMBER: 60/253,625
PRIOR FILING DATE: 2000-11-27
PRIOR APPLICATION NUMBER: 60/257,931
PRIOR FILING DATE: 2000-12-22
PRIOR APPLICATION NUMBER: 60/267,636
PRIOR FILING DATE: 2001-02-09
PRIOR APPLICATION NUMBER: 60/269,308
PRIOR FILING DATE: 2001-02-16

```
  Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
;
US-10-282-122A-41310
; SEQ ID NO 41310
; LENGTH: 2862
; TYPE: DNA
; ORGANISM: Vibrio cholerae

Query Match             66.0%;     Score 19.8;   DB 7;   Length 2862;
Best Local Similarity   87.0%;     Pred. No. 54;
Matches   20;  Conservative    1;  Mismatches    2;  Indels   0;  Gaps   0;

Qy      7 ACCGGTAGAGCAUCATGATCCUG 29
          ||||||  |||:||:||||||||
Db     98 ACCGGTACAGCATCATGATCCCG 120

RESULT 10
US-11-097-143-16217
; Sequence 16217, Application US/11097143
; Publication No. US2005020855BA1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT:    et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION:   ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE
; TITLE OF INVENTION:   DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE:   2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16217
; LENGTH: 2349
; TYPE: DNA
; ORGANISM: DROSOPHILA
US-11-097-143-16217

Query Match             65.3%;     Score 19.6;   DB 10;   Length 2349;
Best Local Similarity   84.6%;     Pred. No. 66;
Matches   22;  Conservative    0;  Mismatches    4;  Indels   0;  Gaps   0;

Qy      1 CAGGCUACGCGTAGAGCAUCATGATC 26
          |||||| ||| || |||||||||||
Db   1851 CAGGCGACGGGTGGAGCACCATGATC 1876

RESULT 11
US-11-097-143-16216/c
; Sequence 16216, Application US/11097143
; Publication No. US2005020855BA1
; GENERAL INFORMATION:
; APPLICANT: Venter, J. Craig
; APPLICANT:    et al.
; TITLE OF INVENTION: DETECTION KIT, SUCH AS NUCLEIC ACID
; TITLE OF INVENTION:   ARRAYS, FOR DETECTING EXPRESSION OF 10,000 OR MORE


; TITLE OF INVENTION:   DROSOPHILA GENES.
; FILE REFERENCE: CL000728
; CURRENT APPLICATION NUMBER: US/11/097,143
; CURRENT FILING DATE:   2005-04-04
; PRIOR APPLICATION NUMBER: 60/157,832
; PRIOR FILING DATE: 1999-10-05
; PRIOR APPLICATION NUMBER: 60/160,191
; PRIOR FILING DATE: 1999-10-19
; PRIOR APPLICATION NUMBER: 60/161,932
; PRIOR FILING DATE: 1999-10-28
; PRIOR APPLICATION NUMBER: 60/164,769
; PRIOR FILING DATE: 1999-11-12
; PRIOR APPLICATION NUMBER: 60/173,383
; PRIOR FILING DATE: 1999-12-28
; PRIOR APPLICATION NUMBER: 60/175,693
; PRIOR FILING DATE: 2000-01-12
; PRIOR APPLICATION NUMBER: 60/184,831
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: 60/191,637
; PRIOR FILING DATE: 2000-03-23
; NUMBER OF SEQ ID NOS: 43008
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 16216
; LENGTH: 4992
; TYPE: DNA
; ORGANISM: DROSOPHILA
US-11-097-143-16216

Query Match             65.3%;     Score 19.6;   DB 10;   Length 4992;
Best Local Similarity   84.6%;     Pred. No. 73;
Matches   22;  Conservative    0;  Mismatches    4;  Indels   0;  Gaps   0;

Qy      1 CAGGCUACGCGTAGAGCAUCATGATC 26
          |||||| ||| || |||||||||||
Db   1561 CAGGCGACGGGTGGAGCACCATGATC 1536

RESULT 12
US-10-425-115-17734/c
; Sequence 17734, Application US/10425115
; Publication No. US20040214272A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION:   Plants
; FILE REFERENCE: 38-21(53222)B
; CURRENT APPLICATION NUMBER: US/10/425,115
; CURRENT FILING DATE:   2003-04-28
; NUMBER OF SEQ ID NOS: 369326
; SEQ ID NO 17734
; LENGTH: 403
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: MRT4577_116177C.1
US-10-425-115-17734

Query Match             64.7%;     Score 19.4;   DB 8;   Length 403;
Best Local Similarity   69.0%;     Pred. No. 66;
Matches   20;  Conservative    3;  Mismatches    6;  Indels   0;  Gaps   0;

Qy      2 AGGCUACGCGTAGAGCAUCATGATCCUGT 30
          ||||| |||| |:|||||| |:|| ||:|
Db    330 AGCCTCCGCGTAGTGGATCAAGATGCTGT 302

RESULT 13
US-10-425-115-139288
; Sequence 139288, Application US/10425115
; Publication No. US20040214272A1
```

```
;   GENERAL INFORMATION:
;   APPLICANT: La Rosa, Thomas J.
;   APPLICANT: Kovalic, David K.
;   APPLICANT: Zhou, Yihua
;   APPLICANT: Cao, Yongwei
;   TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;   TITLE OF INVENTION: Plants
;   FILE REFERENCE: 38-21(53222)B
;   CURRENT APPLICATION NUMBER: US/10/425,115
;   CURRENT FILING DATE: 2003-04-28
;   NUMBER OF SEQ ID NOS: 369326
;   SEQ ID NO 139288
;   LENGTH: 1335
;   TYPE: DNA
;   ORGANISM: Zea mays
;   FEATURE:
;   OTHER INFORMATION: Clone ID: MRT4577_58511C.1
US-10-425-115-139288

Query Match            64.7%;  Score 19.4; DB 8; Length 1335;
Best Local Similarity  72.4%;  Pred. No. 76;
Matches    21; Conservative  2; Mismatches  6; Indels   0; Gaps   0;

Qy     1 CAGGCUACGCGTAGAGCAUCATGATCCUG 29
           ||||||| || |||||| |: ||||| |:|
Db   117 CAGACAACGCGTCGAGGATTATGATGCTG 145


RESULT 14
US-10-767-701-29066
;   Sequence 29066, Application US/10767701
;   Publication No. US2004017268441
;   GENERAL INFORMATION:
;   APPLICANT: Kovalic, David K.
;   APPLICANT: Zhou, Yihua
;   APPLICANT: Cao, Yongwei
;   TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
;   TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
;   FILE REFERENCE: 38-21(53535)B
;   CURRENT APPLICATION NUMBER: US/10/767,701
;   CURRENT FILING DATE: 2004-01-29
;   NUMBER OF SEQ ID NOS: 63128
;   SEQ ID NO 29066
;   LENGTH: 466
;   TYPE: DNA
;   ORGANISM: Sorghum bicolor
;   FEATURE:
;   OTHER INFORMATION: Clone ID: 8548521
US-10-767-701-29066

Query Match            64.0%;  Score 19.2; DB 7; Length 466;
Best Local Similarity  83.3%;  Pred. No. 83;
Matches    20; Conservative  1; Mismatches  1; Indels   3; Gaps   0;

Qy     2 AGGCUACGCGTAGAGCAUCATGAT 25
           |||||:| || |||||||| |||||
Db   301 AGGCTAAGCTTAGAGCAGCATGAT 324


RESULT 15
US-10-762-107-27/c
;   Sequence 27, Application US/10762107
;   Publication No. US2005004329741
;   GENERAL INFORMATION:
;   APPLICANT: ECOPIA BIOSCIENCES INC.
;   APPLICANT: Farnet, Chris
;   APPLICANT: McAlpine, James
;   APPLICANT: Zazopoulos, Emmanuel
;   APPLICANT: Bachmann, Brian
;   APPLICANT: Piraee, Mahmood
;   TITLE OF INVENTION: FARNESYL DIBENZODIAZEPINONE, PROCESSES FOR ITS PRODUCTION AND ITS
;   TITLE OF INVENTION: A PHARMACEUTICAL
;   FILE REFERENCE: 3005-5US
```

```
;   CURRENT APPLICATION NUMBER: US/10/762,107
;   CURRENT FILING DATE: 2004-01-20
;   PRIOR APPLICATION NUMBER: USSN 60/441,126
;   PRIOR FILING DATE: 2003-01-21
;   PRIOR APPLICATION NUMBER: USSN 60/492,997
;   PRIOR FILING DATE: 2003-08-07
;   PRIOR APPLICATION NUMBER: USSN 60/518,286
;   PRIOR FILING DATE: 2003-11-10
;   NUMBER OF SEQ ID NOS: 89
;   SOFTWARE: PatentIn version 3.0
;   SEQ ID NO 27
;   LENGTH: 663
;   TYPE: DNA
;   ORGANISM: Micromonospora sp. strain 046-ECO11
US-10-762-107-27

Query Match            63.3%;  Score 19; DB 8; Length 663;
Best Local Similarity  74.1%;  Pred. No. 1.1e+02;
Matches    20; Conservative  2; Mismatches  5; Indels   0; Gaps   0;

Qy     3 GGCUACGCGTAGAGCAUCATGATCCUG 29
           |||: ||| ||||| ||||| ||||||:|
Db    32 GGCTGGTCGTCGAGCAGCATGATCCTG 6


Search completed: January  8, 2006, 17:18:07
Job time : 563 secs
```

```
                    GenCore version 5.1.6
              Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:        January  8, 2006, 15:55:54 ; Search time 243 Seconds
                                           (without alignments)
                                           89.992 Million cell updates/sec

Title:         US-10-791-367-1
Perfect score: 30
Sequence:      1 caggcuacgcgtagagcaucatgatccugt 30

Scoring table: IDENTITY_NUC
               Gapop 10.0 , Gapext 1.0

Searched:      4637609 seqs, 36446668 residues

Total number of hits satisfying chosen parameters:     9275218

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database  :    Published Applications NA New:*
               1:  /cgn2_6/ptcodata/2/pubpna/US08_NEW_PUB.seq:*
               2:  /cgn2_6/ptcodata/2/pubpna/US06_NEW_PUB.seq:*
               3:  /cgn2_6/ptcodata/2/pubpna/US07_NEW_PUB.seq:*
               4:  /cgn2_6/ptcodata/2/pubpna/PCT_NEW_PUB.seq:*
               5:  /cgn2_6/ptcodata/2/pubpna/US09_NEW_PUB.seq:*
               6:  /cgn2_6/ptcodata/2/pubpna/US10_NEW_PUB.seq:*
               7:  /cgn2_6/ptcodata/2/pubpna/US11_NEW_PUB.seq:*
               8:  /cgn2_6/ptcodata/2/pubpna/US11_NEW_PUB.seq2:*
               9:  /cgn2_6/ptcodata/2/pubpna/US11_NEW_PUB.seq3:*
               10: /cgn2_6/ptcodata/2/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                              SUMMARIES
```

Result No.	Score	Query Match	Length	DB	ID	Description
c 1	18.4	61.3	151169	7	US-11-121-086-38	Sequence 38, Appl
2	17.8	59.3	1110	6	US-10-750-185-25891	Sequence 25891, A
3	17.8	59.3	1110	6	US-10-750-623-25891	Sequence 25891, A
4	17.8	59.3	146656	7	US-11-121-086-68	Sequence 68, Appl
5	17.4	58.0	762	6	US-10-750-185-53338	Sequence 53338, A
c 6	17.4	58.0	762	6	US-10-750-623-53338	Sequence 53338, A
7	17.4	58.0	1002	7	US-11-127-877-20	Sequence 20, Appl
8	17.4	58.0	1733	6	US-10-750-185-26283	Sequence 26283, A
9	17.4	58.0	1733	6	US-10-750-623-26283	Sequence 26283, A
c 10	17.2	57.3	1400	7	US-11-136-527-4382	Sequence 4382, Ap
c 11	17.2	57.3	1696	7	US-11-136-527-224	Sequence 224, App
c 12	17.2	57.3	4347	7	US-11-136-527-286	Sequence 286, App
c 13	17	56.7	201	6	US-10-995-561-11799	Sequence 11799, A
c 14	17	56.7	201	6	US-10-995-561-11837	Sequence 11837, A
c 15	17	56.7	201	6	US-10-995-561-61360	Sequence 61360, A
c 16	17	56.7	4947	6	US-10-995-561-457	Sequence 457, App
c 17	17	56.7	5130	7	US-11-135-855-12	Sequence 12, Appl
c 18	17	56.7	5158	7	US-11-135-855-13	Sequence 13, Appl
c 19	17	56.7	6719	6	US-10-995-561-456	Sequence 456, App
c 20	17	56.7	32157	6	US-10-995-561-13352	Sequence 13352, A
c 21	16.8	56.0	189	7	US-11-108-172-103	Sequence 103, App
c 22	16.8	56.0	261	7	US-11-108-172-749	Sequence 749, App
c 23	16.8	56.0	387	6	US-10-467-657-2599	Sequence 2599, Ap
c 24	16.8	56.0	432	6	US-10-467-657-2601	Sequence 2601, Ap
c 25	16.8	56.0	538	7	US-11-108-172-3	Sequence 3, Appli
c 26	16.8	56.0	2149	6	US-10-880-847-1	Sequence 1, Appli
27	16.8	56.0	28567	7	US-10-947-249-82	Sequence 82, Appl
28	16.8	56.0	179597	7	US-11-121-086-91	Sequence 91, Appl
29	16.6	55.3	335	7	US-11-000-688-871	Sequence 871, App
30	16.6	55.3	260209	6	US-10-993-025-23	Sequence 23, Appl
c 31	16.4	54.7	943	6	US-10-750-185-57644	Sequence 57644, A
c 32	16.4	54.7	943	6	US-10-750-623-57644	Sequence 57644, A
33	16.4	54.7	1670	7	US-11-136-527-2256	Sequence 2256, Ap
34	16.4	54.7	6243	6	US-10-821-234-788	Sequence 788, App
35	16.2	54.0	857	6	US-10-750-185-47478	Sequence 47478, A
36	16.2	54.0	857	6	US-10-750-623-47478	Sequence 47478, A
c 37	16.2	54.0	1669	6	US-10-750-185-29873	Sequence 29873, A
c 38	16.2	54.0	1669	6	US-10-750-623-29873	Sequence 29873, A
39	16.2	54.0	1795	6	US-10-750-185-39532	Sequence 39532, A
40	16.2	54.0	1795	6	US-10-750-623-39532	Sequence 39532, A
41	16.2	54.0	2699	7	US-11-127-877-5	Sequence 5, Appli
c 42	16.2	54.0	4178	6	US-10-964-313-13	Sequence 13, Appl
c 43	16.2	54.0	4473	6	US-10-750-185-53708	Sequence 53708, A
c 44	16.2	54.0	4473	6	US-10-750-623-53708	Sequence 53708, A
c 45	16.2	54.0	37823	7	US-11-124-368A-2880	Sequence 2880, Ap

```
                              ALIGNMENTS

RESULT 1
US-11-121-086-38/c
; Sequence 38, Application US/1121086
; Publication No. US2005026459A1
; GENERAL INFORMATION:
;   APPLICANT: POULSEN, TIM S.
;   APPLICANT: NIELSEN, KIRSTEN V.
;   TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
;   FILE REFERENCE: 09118.6000-00000
;   CURRENT APPLICATION NUMBER: US/11/121,086
;   CURRENT FILING DATE: 2005-05-04
;   PRIOR APPLICATION NUMBER: 60/567,570
;   PRIOR FILING DATE: 2004-05-04
;   NUMBER OF SEQ ID NOS: 107
;   SOFTWARE: PatentIn version 3.3
;   SEQ ID NO 38
;   LENGTH: 151169
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-121-086-38

  Query Match              61.3%;    Score 18.4;   DB 7;   Length 151169;
  Best Local Similarity    67.9%;    Pred. No. 73;
  Matches   19;  Conservative  3;  Mismatches  6;  Indels  0;  Gaps  0;

Qy      1 CAGGCUACGCGTAGAGCAUCATGATCCU 28
            ||| :|| |   || |||||:||||||||:
Db  46374 CAGGATAGGAATATAGCTTCATGATCCT 46347

RESULT 2
US-10-750-185-25891
; Sequence 25891, Application US/10750185
; Publication No. US2005026603A1
; GENERAL INFORMATION:
;   APPLICANT: MMI GENOMICS, INC.
;   APPLICANT: DeNISE, Sue K.
;   APPLICANT: KERR, Richard
;   APPLICANT: ROSENFELD, David
;   APPLICANT: HOLM, Tom
;   APPLICANT: BATES, Stephen
;   APPLICANT: FANTIN, Dennis
;   TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
;   FILE REFERENCE: MMI1100-2
;   CURRENT APPLICATION NUMBER: US/10/750,185
;   CURRENT FILING DATE: 2003-12-31
```

; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25891
; LENGTH: 1110
; TYPE: DNA
; ORGANISM: Bovine 1986681126825
US-10-750-185-25891

Query Match   59.3%; Score 17.8; DB 6; Length 1110;
Best Local Similarity 65.5%; Pred. No. 65;
Matches   19; Conservative   3; Mismatches   3; Indels   7; Gaps   0;

Qy   2 AGGCUACGCGTAGAGCAUCATGATCCUGT 30
              ||| :|| | |||| :||||||||:||
Db   595 AGGGTACACATTCAGCATACTGATCCTGT 623

RESULT 3
US-10-750-623-25891
; Sequence 25891, Application US/10750623
; Publication No. US2005028753lA1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DeNISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MMI1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25891
; LENGTH: 1110
; TYPE: DNA
; ORGANISM: Bovine 1986681126825
US-10-750-623-25891

Query Match   59.3%; Score 17.8; DB 6; Length 1110;
Best Local Similarity 65.5%; Pred. No. 65;
Matches   19; Conservative   3; Mismatches   3; Indels   7; Gaps   0;

Qy   2 AGGCUACGCGTAGAGCAUCATGATCCUGT 30
              ||| :|| | |||| :||||||||:||
Db   595 AGGGTACACATTCAGCATACTGATCCTGT 623

RESULT 4
US-11-121-086-68/c
; Sequence 68, Application US/1121086
; Publication No. US2005026645A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138.6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 68
; LENGTH: 146656
; TYPE: DNA
; ORGANISM: Homo sapiens

US-11-121-086-68

Query Match   59.3%; Score 17.8; DB 7; Length 146656;
Best Local Similarity 65.5%; Pred. No. 1.4e+02;
Matches   19; Conservative   3; Mismatches   3; Indels   7; Gaps   0;

Qy   2 AGGCUACGCGTAGAGCAUCATGATCCUGT 30
              ||| :|| | |||||||||||:||:||
Db   90958 AGACTATAGGTAGAGCATATTGCTCCTGT 90930

RESULT 5
US-10-750-185-53338/c
; Sequence 53338, Application US/10750185
; Publication No. US2005026060A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DeNISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Stephen
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MMI1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 53338
; LENGTH: 762
; TYPE: DNA
; ORGANISM: Bovine 1986681728240
US-10-750-185-53338

Query Match   58.0%; Score 17.4; DB 6; Length 762;
Best Local Similarity 66.7%; Pred. No. 94;
Matches   18; Conservative   3; Mismatches   3; Indels   6; Gaps   0;

Qy   3 GGCUACGCGTAGAGCAUCATGATCCUG 29
              |||:||||| || ||:|| |||||| :|
Db   123 GGCTACTCGGAGTGCATCGTGACCGTG 97

RESULT 6
US-10-750-623-53338/c
; Sequence 53338, Application US/10750623
; Publication No. US2005028753lA1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DeNISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MMI1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 53338
; LENGTH: 762
; TYPE: DNA
; ORGANISM: Bovine 1986681728240
US-10-750-623-53338

Query Match   58.0%; Score 17.4; DB 6; Length 762;

```
Best Local Similarity   66.7%;   Pred. No. 94;
Matches   18;   Conservative   3;   Mismatches   6;   Indels   0;   Gaps   0;

Qy        3 GGCUACGCGTAGAGCAUCAUGATCCUG 29
            ||||:||    |||||| ||||||| ||:|
Db      123 GGCTACTCGGAGTGCATCGTGACCGTG 97

RESULT 7
US-11-127-877-20
; Sequence 20, Application US/11127877
; Publication No. US20050287565A1
; GENERAL INFORMATION:
;   APPLICANT: Merchiers, Pascal G.
;   APPLICANT: Hoffmann, Marcel
;   APPLICANT: Spittaels, Koenraad F. F.
;   APPLICANT: Laenen, Wendy
;   TITLE OF INVENTION: Methods, Compositions and Compound Assays For Inhibiting
;   TITLE OF INVENTION: Amyloid-Beta Protein Production
;   FILE REFERENCE: P27,800-B USA
;   CURRENT APPLICATION NUMBER: US/11/127,877
;   CURRENT FILING DATE: 2005-05-12
;   PRIOR APPLICATION NUMBER: 60/570,352
;   PRIOR FILING DATE: 2004-05-12
;   PRIOR APPLICATION NUMBER: 60/603,948
;   PRIOR FILING DATE: 2004-08-24
;   NUMBER OF SEQ ID NOS: 590
;   SOFTWARE: PatentIn version 3.3
; SEQ ID NO 20
;   LENGTH: 1002
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-11-127-877-20

Query Match              58.0%;   Score 17.4;   DB 7;   Length 1002;
Best Local Similarity   66.7%;   Pred. No. 98;
Matches   18;   Conservative   3;   Mismatches   6;   Indels   0;   Gaps   0;

Qy        4 GCUACGCGTAGAGCAUCATGATCCUGT 30
            ||:||| ||||  ||||| ||||||
Db      611 GCTACGTGGAGATCCTCAGGACCCTGT 637

RESULT 8
US-10-750-185-26283
; Sequence 26283, Application US/10750185
; Publication No. US20050260603A1
; GENERAL INFORMATION:
;   APPLICANT: MMI GENOMICS, INC.
;   APPLICANT: DeNISE, Sue K.
;   APPLICANT: KERR, Richard
;   APPLICANT: ROSENFELD, David
;   APPLICANT: HOLM, Tom
;   APPLICANT: BATES, Stephen
;   APPLICANT: FANTIN, Dennis
;   TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
;   FILE REFERENCE: MMI1100-2
;   CURRENT APPLICATION NUMBER: US/10/750,185
;   CURRENT FILING DATE: 2003-12-31
;   PRIOR APPLICATION NUMBER: US 60/437,482
;   PRIOR FILING DATE: 2002-12-31
;   NUMBER OF SEQ ID NOS: 64922
;   SOFTWARE: PatentIn version 3.1
; SEQ ID NO 26283
;   LENGTH: 1733
;   TYPE: DNA
;   ORGANISM: Bovine 1986881128530
US-10-750-185-26283

Query Match              58.0%;   Score 17.4;   DB 6;   Length 1733;
Best Local Similarity   70.4%;   Pred. No. 1.1e+02;
Matches   19;   Conservative   2;   Mismatches   6;   Indels   0;   Gaps   0;

Qy        4 GCUACGCGTAGAGCAUCATGATCCUGT 30
            ||:|||||     ||||||||||||||:|
Db     1042 GTTACACCAAGAGCAGCATGATCCTTT 1068

RESULT 9
US-10-750-623-26283
; Sequence 26283, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
;   APPLICANT: MMI GENOMICS, INC.
;   APPLICANT: DeNISE, Sue K.
;   APPLICANT: KERR, Richard
;   APPLICANT: ROSENFELD, David
;   APPLICANT: HOLM, Tom
;   APPLICANT: BATES, Stephen
;   APPLICANT: FANTIN, Dennis
;   TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
;   FILE REFERENCE: MMI1100-1
;   CURRENT APPLICATION NUMBER: US/10/750,623
;   CURRENT FILING DATE: 2003-12-31
;   PRIOR APPLICATION NUMBER: US 60/437,482
;   PRIOR FILING DATE: 2002-12-31
;   NUMBER OF SEQ ID NOS: 64922
;   SOFTWARE: PatentIn version 3.1
; SEQ ID NO 26283
;   LENGTH: 1733
;   TYPE: DNA
;   ORGANISM: Bovine 1986881128530
US-10-750-623-26283

Query Match              58.0%;   Score 17.4;   DB 6;   Length 1733;
Best Local Similarity   70.4%;   Pred. No. 1.1e+02;
Matches   19;   Conservative   2;   Mismatches   6;   Indels   0;   Gaps   0;

Qy        4 GCUACGCGTAGAGCAUCATGATCCUGT 30
            ||:|||||     ||||||||||||||:|
Db     1042 GTTACACCAAGAGCAGCATGATCCTTT 1068

RESULT 10
US-11-136-527-4382/c
; Sequence 4382, Application US/11136527
; Publication No. US20050287570A1
; GENERAL INFORMATION:
;   APPLICANT: Wyeth
;   APPLICANT: Mounts, William M
;   TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
;   FILE REFERENCE: 031896-041000 (AM101086)
;   CURRENT APPLICATION NUMBER: US/11/136,527
;   CURRENT FILING DATE: 2005-05-25
;   PRIOR APPLICATION NUMBER: US 60/574,294
;   PRIOR FILING DATE: 2005-05-26
;   NUMBER OF SEQ ID NOS: 362830
;   SOFTWARE: PatentIn version 3.2
; SEQ ID NO 4382
;   LENGTH: 1400
;   TYPE: DNA
;   ORGANISM: Rattus norvegicus
US-11-136-527-4382

Query Match              57.3%;   Score 17.2;   DB 7;   Length 1400;
Best Local Similarity   63.3%;   Pred. No. 1.3e+02;
Matches   19;   Conservative   3;   Mismatches   8;   Indels   0;   Gaps   0;

Qy        1 CAGGCUACGCGTAGAGCAUCATGATCCUGT 30
            |||| :||| |||||| |||:|  |||||| ||
Db      511 CAGCATCCGTGTAGGGCATCTGGCTCCTGT 482

RESULT 11
US-11-136-527-224/c
; Sequence 224, Application US/1136527
```

; Publication No. US2005028570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 224
; LENGTH: 1696
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-224

Query Match   57.3%; Score 17.2; DB 7; Length 1696;
Best Local Similarity   66.7%; Pred. No. 1.3e+02;
Matches 20; Conservative 2; Mismatches 8; Indels 0; Gaps 0;

Qy   1 CAGGCUACGCGTAGAGCAUCATGATCCUGT 30
            |||||:|||||:||| ||||:||||
Db   1063 CAGGCTCCCGTGTAGGGCATCTTCATTCGGT 1034

RESULT 12
US-11-136-527-286/c
; Sequence 286, Application US2005028570A1
; Publication No. US2005028570A1
; GENERAL INFORMATION:
; APPLICANT: Wyeth
; APPLICANT: Mounts, William M
; TITLE OF INVENTION: Probe Arrays For Expression Profiling of Rat Genes
; FILE REFERENCE: 031896-041000 (AM101086)
; CURRENT APPLICATION NUMBER: US/11/136,527
; CURRENT FILING DATE: 2005-05-25
; PRIOR APPLICATION NUMBER: US 60/574,294
; PRIOR FILING DATE: 2005-05-26
; NUMBER OF SEQ ID NOS: 362830
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 286
; LENGTH: 4347
; TYPE: DNA
; ORGANISM: Rattus norvegicus
US-11-136-527-286

Query Match   57.3%; Score 17.2; DB 7; Length 4347;
Best Local Similarity   63.3%; Pred. No. 1.5e+02;
Matches 19; Conservative 3; Mismatches 8; Indels 0; Gaps 0;

Qy   1 CAGGCUACGCGTAGAGCAUCATGATCCUGT 30
            ||||:: |||||:||||||:||||
Db   3458 CAGCTCCCGTGTAGGGCATCTGGCTCCTGT 3429

RESULT 13
US-10-995-561-11799/c
; Sequence 11799, Application US/10995561
; Publication No. US2005027205A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11799
; LENGTH: 201

; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-11799

Query Match   56.7%; Score 17; DB 6; Length 201;
Best Local Similarity   72.0%; Pred. No. 1.2e+02;
Matches 18; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy   1 CAGGCUACGCGTAGAGCAUCATGAT 25
            |||||:| |||| ||:|||||
Db   179 CAGGCTCCACGAAGACCATCATGGT 155

RESULT 14
US-10-995-561-11837/c
; Sequence 11837, Application US/10995561
; Publication No. US2005027205A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 11837
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-11837

Query Match   56.7%; Score 17; DB 6; Length 201;
Best Local Similarity   72.0%; Pred. No. 1.2e+02;
Matches 18; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy   1 CAGGCUACGCGTAGAGCAUCATGAT 25
            |||||:| |||| ||:|||||
Db   179 CAGGCTCCACGAAGACCATCATGGT 155

RESULT 15
US-10-995-561-61360/c
; Sequence 61360, Application US/10995561
; Publication No. US2005027205A1
; GENERAL INFORMATION:
; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE, METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 61360
; LENGTH: 201
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-995-561-61360

Query Match   56.7%; Score 17; DB 6; Length 201;
Best Local Similarity   72.0%; Pred. No. 1.2e+02;
Matches 18; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

Qy   1 CAGGCUACGCGTAGAGCAUCATGAT 25
            |||||:| |||| ||:|||||
Db   176 CAGGCTCCACGAAGACCATCATGGT 152

Search completed: January 8, 2006, 17:22:16
Job time : 245 secs

THIS PAGE BLANK (USPTO)

```
                    GenCore version 5.1.6
              Copyright (c) 1993 - 2006  Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:        January 8, 2006, 15:02:49 ; Search time 1726 Seconds
                                          (without alignments)
                                          988.009 Million cell updates/sec

Title:         US-10-791-367-1
Perfect score: 30
Sequence:      1 caggcuacgcgtagagcaucatgatccugt 30

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched:      5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters:    11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :     GenEmbl:*
               1: gb_ba:*
               2: gb_in:*
               3: gb_env:*
               4: gb_om:*
               5: gb_ov:*
               6: gb_pat:*
               7: gb_ph:*
               8: gb_pr:*
               9: gb_ro:*
               10: gb_sts:*
               11: gb_sy:*
               12: gb_un:*
               13: gb_vi:*
               14: gb_htg:*
               15: gb_pl:*
```

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	30	100.0	30	6	CS079386	CS079386 Sequence
2	30	100.0	30	6	AR481874	AR481874 Sequence
3	30	100.0	31	6	BD249885	BD249885 Aptamers
4	30	100.0	31	6	AR366193	AR366193 Sequence
5	30	100.0	31	6	AR442638	AR442638 Sequence
c 6	21.6	72.0	214332	14	AC120945	AC120945 Rattus no
7	21	70.0	226789	14	AC123112	AC123112 Rattus no
8	20	66.7	6002	15	AJ616913	AJ616913 Sordaria
9	20	66.7	110000	1	BX571965_10	Continuation (11 o
c 10	20	66.7	110000	1	CP000010_23	Continuation (24 o
11	20	66.7	163878	2	AC150398	AC150398 Branchios
c 12	20	66.7	172391	14	AC155516	AC155516 Zea mays
13	20	66.7	174912	14	AC145730	AC145730 Zea mays
c 14	20	66.7	181826	14	AC155574	AC155574 Zea mays
15	19.8	66.0	10205	1	AE004320	AE004320 Vibrio ch
16	19.6	65.3	606	6	AR506387	AR506387 Sequence
17	19.6	65.3	615	6	AR509529	AR509529 Sequence
18	19.6	65.3	705	6	AR505666	AR505666 Sequence
19	19.6	65.3	2186	2	AY058696	AY058696 Drosophil
20	19.6	65.3	2349	6	CQ588459	CQ588459 Sequence
21	19.6	65.3	2782	6	AY052083	AY052083 Drosophil
22	19.6	65.3	4992	6	CQ588458	CQ588458 Sequence
23	19.6	65.3	49054	14	AC019890	AC019890 Drosophil
24	19.6	65.3	104135	8	AC113373	AC113373 Homo sapi
25	19.6	65.3	123275	8	AC008779	AC008779 Homo sapi
26	19.6	65.3	175867	8	AC007976	AC007976 Drosophil
c 27	19.6	65.3	185689	13	AY261364	AY261364 African s
28	19.6	65.3	346734	2	AE003584	AE003584 Drosophil
29	19.4	64.7	153129	9	AC163209	AC163209 Mus muscu
c 30	19.4	64.7	169815	9	AC122057	AC122057 Mus muscu
31	19.4	64.7	175651	14	AC150605	AC150605 Callithri
c 32	19.4	64.7	230041	9	AC111047	AC111047 Mus muscu
33	19.4	64.7	251257	14	AC095972	AC095972 Rattus no
34	19.2	64.0	159866	8	AC016598	AC016598 Homo sapi
35	19	63.3	603	6	AR385986	AR385986 Sequence
36	19	63.3	663	6	CQ879025	CQ879025 Sequence
37	19	63.3	1269	6	AX653992	AX653992 Sequence
38	19	63.3	1692	5	BC085826	BC085826 Rattus no
c 39	19	63.3	36602	6	CQ878999	CQ878999 Sequence
c 40	19	63.3	44343	1	AP004312	AP004312 Synechocy
41	19	63.3	87792	15	AP003837	AP003837 Oryza sat
42	19	63.3	110000	1	AE014295_21	Continuation (22 o
c 43	19	63.3	110000	15	AP008215_058	Continuation (59 o
c 44	19	63.3	114104	9	BX545849	BX545849 Mouse DNA
c 45	19	63.3	151721	15	AP003943	AP003943 Oryza sat

ALIGNMENTS

```
RESULT 1
CS079386
LOCUS       CS079386                       30 bp    DNA            linear   PAT 06-MAY-2005
DEFINITION  Sequence 21 from Patent WO2005020972.
ACCESSION   CS079386
VERSION     CS079386.1  GI:63093789
KEYWORDS    .
SOURCE      synthetic construct
  ORGANISM  synthetic construct
            other sequences; artificial sequences.
REFERENCE   1
  AUTHORS   Shima,D., Calias,P. and Adamis,A.P.
  TITLE     Combination therapy for the treatment of ocular neovascular
            disorders
  JOURNAL   Patent: WO 2005020972-A 21 10-MAR-2005;
            Eyetech Pharmaceuticals, Inc. (US)
FEATURES             Location/Qualifiers
     source          1..30
                     /organism="synthetic construct"
                     /mol_type="unassigned DNA"
                     /db_xref="taxon:32630"
                     /note="Description of Artificial Sequence: Synthetic
                     anti-PDGF aptamer"
ORIGIN

Query Match             100.0%;  Score 30;  DB 6;  Length 30;
  Best Local Similarity   90.0%;  Pred. No. 0.0057;
  Matches   27;  Conservative    3;  Mismatches    0;  Indels    0;  Gaps    0;

Qy          1 CAGGCUACGCGTAGAGCAUCATGATCCUGT 30
              |||||||||||||||||||||||:|||:||:
Db          1 CAGGCTACGCGTAGAGCATCATGATCCTGT 30


RESULT 2
AR481874
LOCUS       AR481874                       30 bp    DNA            linear   PAT 14-MAY-2004
DEFINITION  Sequence 1 from patent US 6699843.
ACCESSION   AR481874
VERSION     AR481874.1  GI:47243605
KEYWORDS    .
```

```
SOURCE       Unknown.
ORGANISM     Unknown.
             Unclassified.
REFERENCE    1  (bases 1 to 30)
AUTHORS      Pietras,K., Ostman,A., Heldin,C.-H. and Rubin,K.
TITLE        Method for treatment of tumors using nucleic acid ligands to PDGF
JOURNAL      Patent: US 6699843-A 1 02-MAR-2004;
             Gilead Sciences, Inc.; Foster City, CA
FEATURES             Location/Qualifiers
     source          1..30
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Query Match             100.0%;  Score 30;  DB 6;  Length 30;
  Best Local Similarity   90.0%;  Pred. No. 0.0057;
  Matches   27;  Conservative    3;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          1 CAGGCUACGCGUAGAGCAUCAUGATGATCCUGT 30
              ||||:|||||||||||||||:||||||:|
Db          1 CAGGCTACGCGTAGAGCATCATGATCCTGT 30


RESULT 3
BD249885
LOCUS       BD249885                  31 bp    DNA     linear   PAT 17-JUL-2003
DEFINITION  Aptamers as reagents for high throughput screening.
ACCESSION   BD249885
VERSION     BD249885.1  GI:33059655
KEYWORDS    JP 2002541434-A/3.
SOURCE      synthetic construct
ORGANISM    synthetic construct
            other sequences; artificial sequences.
REFERENCE    1  (bases 1 to 31)
AUTHORS      Janjic,N. and Gold,L.
TITLE        Aptamers as reagents for high throughput screening
JOURNAL      Patent: JP 2002541434-A 3 03-DEC-2002;
             NEXSTAR PHARMACEUTICALS INC
COMMENT     OS   Artificial Sequence
            PN   JP 2002541434-A/3
            PD   03-DEC-2002
            PF   01-FEB-2000 JP 2000598668
            PR   09-FEB-1999 US 09/246461
            PI   NEBOJSA JANJIC,LARRY GOLD
            PC   G01N33/566,C12Q1/68,G01N33/15,G01N33/50,G01N37/00 CC
                 G01N33/566,C12Q1/68,G01N33/15,G01N33/53,G01N37/00 CC The
            Description of Artificial Sequence: Nucleic acid ligand CC The
            t at position 1 is 2'-NH2.
            CC   The u's and c's at positions 7, 9, 20, 21, 27, 28 and 29 are
            CC   2'-F.
            CC   The a's and g's at positions 10, 15, 17, 22 and 30 are 2'-O-
            CC   methyl.
            CC   The g at position 10 and the a at position 22 have been CC
                 modified with
            CC   polyethylene glycol of 18 units.
            FH   Key             Location/Qualifiers
            FT   modified_base   (1)..(31)
            FT   modified_base   (1)..(31).
FEATURES             Location/Qualifiers
     source          1..31
                     /organism="synthetic construct"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:32630"
ORIGIN

Query Match             100.0%;  Score 30;  DB 6;  Length 31;
  Best Local Similarity   90.0%;  Pred. No. 0.0057;
  Matches   27;  Conservative    3;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          1 CAGGCUACGCGUAGAGCAUCAUGATGATCCUGT 30
              ||||:|||||||||||||||:||||||:|
Db          1 CAGGCTACGCGTAGAGCATCATGATCCTGT 30


RESULT 4
AR366193
LOCUS       AR366193                  31 bp    DNA     linear   PAT 12-SEP-2003
DEFINITION  Sequence 3 from patent US 6329145.
ACCESSION   AR366193
VERSION     AR366193.1  GI:34598575
KEYWORDS    .
SOURCE      Unknown.
ORGANISM    Unknown.
            Unclassified.
REFERENCE    1  (bases 1 to 31)
AUTHORS      Janjic,N. and Gold,L.
TITLE        Determining non-nucleic acid molecule binding to target by
             competition with nucleic acid ligand
JOURNAL      Patent: US 6329145-A 3 11-DEC-2001;
             Gilead Science, Inc.; Foster City, CA
FEATURES             Location/Qualifiers
     source          1..31
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Query Match             100.0%;  Score 30;  DB 6;  Length 31;
  Best Local Similarity   90.0%;  Pred. No. 0.0057;
  Matches   27;  Conservative    3;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          1 CAGGCUACGCGUAGAGCAUCAUGATGATCCUGT 30
              ||||:|||||||||||||||:||||||:|
Db          2 CAGGCTACGCGTAGAGCATCATGATCCTGT 31


RESULT 5
AR442638
LOCUS       AR442638                  31 bp    DNA     linear   PAT 20-FEB-2004
DEFINITION  Sequence 3 from patent US 6670132.
ACCESSION   AR442638
VERSION     AR442638.1  GI:42669895
KEYWORDS    .
SOURCE      Unknown.
ORGANISM    Unknown.
            Unclassified.
REFERENCE    1  (bases 1 to 31)
AUTHORS      Janjic,N. and Gold,L.
TITLE        Determining non-nucleic acid molecule binding to target by
             competition with nucleic acid ligand
JOURNAL      Patent: US 6670132-A 3 30-DEC-2003;
             Gilead Sciences, Inc.; Foster City, CA
FEATURES             Location/Qualifiers
     source          1..31
                     /organism="unknown"
                     /mol_type="genomic DNA"
ORIGIN

Query Match             100.0%;  Score 30;  DB 6;  Length 31;
  Best Local Similarity   90.0%;  Pred. No. 0.0057;
  Matches   27;  Conservative    3;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          1 CAGGCUACGCGUAGAGCAUCAUGATGATCCUGT 30
              ||||:|||||||||||||||:||||||:|
Db          2 CAGGCTACGCGTAGAGCATCATGATCCTGT 31


RESULT 6
AC120945/c
LOCUS       AC120945              214332 bp    DNA     linear   HTG 19-NOV-2002
DEFINITION  Rattus norvegicus clone CH230-285A15, WORKING DRAFT SEQUENCE.
ACCESSION   AC120945
VERSION     AC120945.7  GI:25091832
KEYWORDS    HTG; HTGS_PHASE2; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Rattus norvegicus (Norway rat)
ORGANISM    Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Query Match             100.0%;  Score 30;  DB 6;  Length 31;
  Best Local Similarity   90.0%;  Pred. No. 0.0057;
  Matches   27;  Conservative    3;  Mismatches   0;  Indels   0;  Gaps   0;

Qy          1 CAGGCUACGCGTAGAGCAUCATGATCCUGT 30
              ||||:|||||||||||||||:||||||:|
Db          2 CAGGCTACGCGTAGAGCATCATGATCCTGT 31
```

REFERENCE   1 (bases 1 to 214332)
AUTHORS   Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
Muzny,D.Marie., Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
Baldwin,S., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
Draper,H., Dugan-Rocha,S., Dunn,A., Durbin,K., Duval,B., Eaves,K.,
Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,K.,
Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
Lorensuhewa,L., Louiseged,H., Lozado,R.J., Lu,X., Ma,J.,
Maheshwari,M., Mahindarne,M., Mahmoud,M., Malloy,K., Mangum,A.,
Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Parks,K.,
Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Parks,K.,
Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
Rives,C., Rodkey,T., Rojas,R., Rose,M., Rose,R., Ruiz,S.J.,
Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
Shetty,J., Shvartsbeyn,A., Sisson,I., Sitter,C.D., Smajs,D.,
Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
Williams,G., Willson,R., Wleczyk,R., Wooden,H., Worley,K.,
Wright,D., Wright,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhao,S., Dunn,D., von
Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
Weinstock,G. and Gibbs,R.A.
TITLE   Direct Submission
JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 214332)
AUTHORS   Worley,K.C.
TITLE   Direct Submission
JOURNAL   Submitted (14-MAY-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3 (bases 1 to 214332)
AUTHORS   Rat Genome Sequencing Consortium.
TITLE   Direct Submission
JOURNAL   Submitted (19-NOV-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
COMMENT   On Nov 19, 2002 this sequence version replaced gi:23907982.
The sequence in this assembly is a combination of BAC based reads
and whole genome shotgun sequencing reads assembled using Atlas
(http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
in the feature table below represents a scaffold in the Atlas
assembly (a 'contig-scaffold'). Within each contig-scaffold,
individual sequence contigs are ordered and oriented, and separated
by sized gaps filled with Ns to the estimated size. The sequence

may extend beyond the ends of the clone and there may be sequence
contigs within a contig-scaffold that consist entirely of whole
genome shotgun sequence reads. Both end sequences and whole genome
shotgun sequence only contigs will be indicated in the feature
table.
------ Genome Center
Center: Baylor College of Medicine
Center code: BCM
Web site: http://www.hgsc.bcm.tmc.edu/
Contact: hgsc-help@bcm.tmc.edu
------ Project Information
Center project name: GWRD
Center clone name: CH230-285A15
------ Summary Statistics
Assembly program: Phrap; version 0.990329
Consensus quality: 192159 bases at least Q40
Consensus quality: 194070 bases at least Q30
Consensus quality: 195346 bases at least Q20
Estimated insert size: 197053; sum-of-contigs estimation
Quality coverage: 9x in Q20 bases; sum-of-contigs estimation
------
* NOTE: Estimated insert size may differ from sequence length
* (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submittor.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
*   1   214332: contig of 214332 bp in length.
FEATURES   Location/Qualifiers
source   1..214332
/organism="Rattus norvegicus"
/mol_type="genomic DNA"
/db_xref="taxon:10116"
/clone="CH230-285A15"
misc_feature   1..1288
/note="wgs_end_extension
clone_end:Sp6"
misc_feature   5411..6442
/note="wgs_end_extension
clone_end:Sp6"
misc_feature   8136..9424
/note="clone_boundary
clone_end:Sp6
site:
end_sequence:BZ259845"
misc_feature   204676..205494
/note="clone_boundary
clone_end:T7
site:
end_sequence:BZ259842"
misc_feature   213050..214332
/note="wgs_end_extension
clone_end:T7"
ORIGIN

Query Match   72.0%;   Score 21.6;   DB 14;   Length 214332;
Best Local Similarity   75.0%;   Pred. No. 99;
Matches   21;   Conservative   3;   Mismatches   4;   Indels   0;   Gaps   0;

Qy   1 CAGGCTCACGCGTGAGCAUCATGATCCU 28
     |||||.|||||||:||||||||:||||||||:|||
Db   144601 CAGGCTGCCCGTGAGCATCATGAGCAT 144574

RESULT 7
AC123112
LOCUS   AC123112   226789 bp   DNA   linear   HTG 13-MAY-2003
DEFINITION   Rattus norvegicus clone CH230-174P11, WORKING DRAFT SEQUENCE, 2

```
ACCESSION   AC123112
VERSION     AC123112.4  GI:30580864
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
SOURCE      Rattus norvegicus (Norway rat)
  ORGANISM  Rattus norvegicus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
            Sciurognathi; Muroidea; Muridae; Murinae; Rattus.
REFERENCE   1  (bases 1 to 226789)
  AUTHORS   Muzny,D.Marie, Metzker,M.Lee., Abramzon,S., Adams,C., Alder,J.,
            Allen,C., Allen,H., Alsbrooks,S., Amin,A., Anguiano,D.,
            Anyalebechi,V., Aoyagi,A., Ayodeji,M., Baca,E., Baden,H.,
            Baldwin,D., Bandaranaike,D., Barber,M., Barnstead,M., Benahmed,F.,
            Biswalo,K., Blair,J., Blankenburg,K., Blyth,P., Brown,M.,
            Bryant,N., Buhay,C., Burch,P., Burrell,K., Calderon,E.,
            Cardenas,V., Carter,K., Cavazos,I., Ceasar,H., Center,A.,
            Chacko,J., Chavez,D., Chen,G., Chen,R., Chen,Y., Chen,Z., Chu,J.,
            Cleveland,C., Cockrell,R., Cox,C., Coyle,M., Cree,A., D'Souza,L.,
            Davila,M.L., Davis,C., Davy-Carroll,L., De Anda,C., Dederich,D.,
            Delgado,O., Denson,S., Deramo,C., Ding,Y., Dinh,H., Divya,K.,
            Draper,H., Dugan-Rocha,S., Dunn,K., Duval,B., Eaves,K.,
            Egan,A., Escotto,M., Eugene,C., Evans,C.A., Falls,T., Fan,G.,
            Fernandez,S., Finley,M., Flagg,N., Forbes,L., Foster,M., Foster,P.,
            Fraser,C.M., Gabisi,A., Ganta,R., Garcia,A., Garner,T., Garza,M.,
            Gebregeorgis,E., Geer,K., Gill,R., Grady,M., Guerra,W., Guevara,W.,
            Gunaratne,P., Haaland,W., Hamil,C., Hamilton,C., Hamilton,J.,
            Harvey,Y., Havlak,P., Hawes,A., Henderson,N., Hernandez,J.,
            Hernandez,R., Hines,S., Hladun,S.L., Hodgson,A., Hogues,M.,
            Hollins,B., Howells,S., Hulyk,S., Hume,J., Idlebird,D., Jackson,A.,
            Jackson,L., Jacob,L., Jiang,H., Johnson,B., Johnson,R., Jolivet,A.,
            Karpathy,S., Kelly,S., Kelly,S., Khan,Z., King,L., Kovar,C.,
            Kowis,C., Kraft,C.L., Lebow,H., Levan,J., Lewis,L., Li,Z., Liu,J.,
            Liu,J., Liu,W., Liu,Y., London,P., Longacre,S., Lopez,J.,
            Lorensuhewa,L., Louiseged,H., Lozado,R.J., Lu,X., Ma,J.,
            Maheshwari,M., Mahindartne,M., Mahmoud,M., Malloy,K., Mangum,A.,
            Mangum,B., Mapua,P., Martin,K., Martin,R., Martinez,E.,
            Mawhiney,S., McLeod,M.P., McNeill,T.Z., Meenen,E.,
            Milosavljevic,A., Miner,G., Minja,E., Montemayor,J., Moore,S.,
            Morgan,M., Morris,K., Morris,S., Munidasa,M., Murphy,M., Nair,L.,
            Nankervis,C., Neal,D., Newton,N., Nguyen,N., Norris,S., Parks,K.,
            Nwaokelemeh,O., Okwuonu,G., Olarnpunsagoon,A., Pal,S., Perez,A.,
            Pasternak,S., Paul,H., Perez,A., Perez,L., Pfannkoch,C.,
            Plopper,F., Poindexter,A., Popovic,D., Primus,E., Pu,L.-L.,
            Puazo,M., Quiroz,J., Rachlin,E., Reeves,K., Regier,M.A., Reigh,R.,
            Reilly,B., Reilly,M., Ren,Y., Reuter,M., Richards,S., Riggs,F.,
            Rives,C., Rodkey,T., Rojas,A., Rose,M., Rose,R., Ruiz,S.J.,
            Sanders,W., Savery,G., Scherer,S., Scott,G., Shatsman,S., Shen,H.,
            Shetty,J., Shvartsbeyn,A., Sisson,J., Sitter,C.D., Smajs,D.,
            Sneed,A., Sodergren,E., Song,X.-Z., Sorelle,R., Sosa,J.,
            Steimle,M., Strong,R., Sutton,A., Svatek,A., Tabor,P., Taylor,C.,
            Taylor,T., Thomas,N., Thomas,S., Tingey,A., Trejos,Z., Usmani,K.,
            Valas,R., Vera,V., Villasana,D., Waldron,L., Walker,B., Wang,J.,
            Wang,Q., Wang,S., Warren,J., Warren,R., Wei,X., White,F.,
            Williams,G., Willson,R., Wu,J., Yakub,S., Yen,J., Yoon,L., Yoon,V.,
            Yu,F., Zhang,J., Zhou,J., Zhou,X., Zhou,S., Dunn,D., von
            Niederhausern,A., Weiss,R., Smith,D.R., Holt,R.A., Smith,H.O.,
            Weinstock,G. and Gibbs,R.A.
  TITLE     Direct Submission
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 226789)
  AUTHORS   Worley,K.C.
  TITLE     Direct Submission
  JOURNAL   Submitted (29-MAY-2002) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
REFERENCE   3  (bases 1 to 226789)
  AUTHORS   Rat Genome Sequencing Consortium.
  TITLE     Direct Submission
  JOURNAL   Submitted (13-MAY-2003) Human Genome Sequencing Center, Department
            of Molecular and Human Genetics, Baylor College of Medicine, One
            Baylor Plaza, Houston, TX 77030, USA
COMMENT     On May 13, 2003 this sequence version replaced gi:23265213.
            The sequence in this assembly is a combination of BAC based reads
            and whole genome shotgun sequencing reads assembled using Atlas
            (http://www.hgsc.bcm.tmc.edu/projects/rat/). Each contig described
            in the feature table below represents a scaffold in the Atlas
            assembly (a 'contig-scaffold'). Within each contig-scaffold,
            individual sequence contigs are ordered and oriented, and separated
            by sized gaps filled with Ns to the estimated size. The sequence
            may extend beyond the ends of the clone and there may be sequence
            contigs within a contig-scaffold that consist entirely of whole
            genome shotgun sequence reads. Both end sequences and whole genome
            shotgun sequence only contigs will be indicated in the feature
            table.
            --------------- Genome Center
            Center: Baylor College of Medicine
            Center code: BCM
            Web site: http://www.hgsc.bcm.tmc.edu
            Contact: hgsc-help@bcm.tmc.edu
            --------------- Project Information
            Center project name: GLAC
            Center clone name: CH230-174P11
            --------------- Summary Statistics
            Assembly program: Atlas 3.0;
            Consensus quality: 219487 bases at least Q40
            Consensus quality: 221101 bases at least Q30
            Consensus quality: 222209 bases at least Q20
            Estimated insert size: 227073; sum-of-contigs estimation
            Quality coverage: 7x in Q20 bases; sum-of-contigs estimation
          * NOTE: Estimated insert size may differ from sequence length
          *   (see http://www.hgsc.bcm.tmc.edu/docs/Genbank_draft_data.html).
          * NOTE: This is a 'working draft' sequence. It currently
          *   consists of 2 contigs. The true order of the pieces
          *   is not known and their order in this sequence record is
          *   arbitrary. Gaps between the contigs are represented as
          *   runs of N, but the exact sizes of the gaps are unknown.
          *   This record will be updated with the finished sequence
          *   as soon as it is available and the accession number will
          *   be preserved.
          *     1  225474: contig of 225474 bp in length
          *     225474: gap of unknown length
          *     225475  225474: gap of unknown length
          *     225575  226789: contig of 1215 bp in length.
FEATURES             Location/Qualifiers
     source          1..226789
                     /organism="Rattus norvegicus"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:10116"
                     /clone="CH230-174P11"
     misc_feature    1..1460
                     /note="wgs end_extension
                     clone_end:T7"
     misc_feature    complement(3044..3574)
                     /note="clone_boundary
                     clone_end:T7
                     site:EcoRI
                     end_sequence:BH350388"
     gap             225475..225574
                     /estimated_length=unknown
ORIGIN
Query Match              70.0%;  Score 21;  DB 14;  Length 226789;
Best Local Similarity    72.4%;  Pred. No. 1.9e+02;
Matches   21;  Conservative   3;  Mismatches   5;  Indels   0;  Gaps   0;

Qy        1 CAGCCUACGCGTAGCAUCATGATCCUG 29
            ||||.:||||||||||.:||||||.:||
Db   187544 CAGGATACCCTTAGAGCATCAGCATCCTG 187572

RESULT 8
AJ616913
LOCUS       AJ616913        6002 bp   DNA    linear   PLN 11-JUN-2004
DEFINITION  Sordaria macrospora pro40 gene for putative signalling cascade
```

```
             component, exons 1-2.
ACCESSION    AJ616913
VERSION      AJ616913.1  GI:39645990
KEYWORDS     pro40 gene; putative signalling cascade component.
SOURCE       Sordaria macrospora
ORGANISM     Sordaria macrospora
             Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
             Sordariomycetidae; Sordariales; Sordariaceae; Sordaria.
REFERENCE    1
AUTHORS      Witzel-Schloemp,K., Strauch,P., Zhang,H. and Kueck,U.
TITLE        Molecular genetic characterization of the developmental mutant
             pro40 from Sordaria macrospora: Identification of a putative
             component of a signaling cascade
JOURNAL      Unpublished
REFERENCE    2 (bases 1 to 6002)
AUTHORS      Witzel-Schloemp,K.
TITLE        Direct Submission
JOURNAL      Submitted (09-DEC-2003) Witzel-Schloemp K., LS Allgemeine und
             Molekulare Botanik, Ruhr-Universitaet Bochum, Universitaetsstrasse
             150, 44780 Bochum, GERMANY
FEATURES     Location/Qualifiers
     source  1..6002
             /organism="Sordaria macrospora"
             /mol_type="genomic DNA"
             /strain="wild type S 48977"
             /db_xref="taxon:5147"
     gene    348..4360
             /gene="pro40"
     mRNA    join(348..3768,3831..4360)
             /gene="pro40"
     CDS     join(348..3768,3831..4360)
             /gene="pro40"
             /codon_start=1
             /product="putative signalling cascade component"
             /protein_id="CAE83713.1"
             /db_xref="GI:39645991"
             /translation="MSRPRGFQDEELFISRPRPSDEDCSPTVEPLRIFKPMSPKPPG
             AADKASYSKYPAPPTSSSAASASASIAPLPGLPGLPPLPSFPLPPGASSSAAPLPYPD
             DDLSSRPTQGASKPARAPYPVDDVRRPSPASSTGRIYSPPPISSTPQLHNDSSPRLNI
             SPIEKKPGLADRRGAVPKPLQSPDSPTGDDLFAKPLGRPVQQPAQQQAQQHPAQRISN
             AYQQKPYYPPPGGGNVRKNSEPSIARISSTASTSTTRASRGSPPPPETPIVEPGVIPG
             GGIEARYAAAGISGTATLTSLQAQTAQSAAANARLAQYGGQRPPSQPPHXQHPALAQS
             QQQPPQAQAQPPRWTPTESPDDQPFGPPTVYQGDAVAQNPQPARPQPQQQNSYNLPK
             NPEPAGSNAPAGSSLQVSVLEQDFQRMQASTPPPAYSSVTPNAASSGYPAEKQRPTQA
             HPATATATATATTNNNNNNNNNNNNNNATPNPTSRPATAASGSSRPANTASPAPSS
             SIHKPKPVAVAAPQTPAQQHPCHPAFANEPQLQHPAQAAHNVGHPALQHTPSLMASSQ
             APPPLPEGWIAHLDQNSGQYYYIHLATQATQWEFPKGPNPISHEAAPLSPTASTYGNP
             LASPLLGGAGLASPMFHPQTPGYAESIMTVASAAPSGFTGPPSAGVDMYKIMPTNG
             VYFGPYLRYVNMDLEKGIWHGSIMIVTDAPQPPTIHIHLSVDLSPNRQLIPHNIWTH
             QRWVFYKYDMDLQMSEHGTERWTYAVTSHLGCTRYEBFIVAGRYETGWRMIAHSGNDFA
             PSTNQNERAKLGGVGFMWKDILQKNVECGGFHVQLGLGDQIYGDRLWREVPILKQWLS
             IAGRENRKNVPWTARHEEDVTHAYPHYTSHFDQPFMREAFAQIPHVLQIDDHDIFDG
             FGSYPEYMQSSAIFKNIGRIAIDMYLLFQHHTTVEMLRNVSSDMLFTITGAGWHFVK
             YLGPAVVVGPDCRSERTQTRVMAGPTYQGLFPKVAMLPPSVQHCIWMLSVPVVYPRL
             ETVETLANTFATGKKAVNSTYNLLGKVTSSVAGVVGGKBVVAHGFKEVKRAVGKSGLM
             GNVLNQFGEPDIAEELKDLWTHESKDLERTYLIRTLQGIAAQKGVRMTFLSGDVSSAG
             AGLLHDPSHPSDHKTMYQLITSPIVAAPCSSYLLKALHSGSNNKLLYVPLNGHKSTHE
             VSDTKEDMMEIFHTDASGAAREYKKLMARRNYVAIVAYDPDAAVNGGMGGSMMNGQGG
             GGYAASVHSGGSGGNGGGLNKLSLAVDFVVQGDGAFTATTKYGPVIVPHLEYGH"
     exon    348..3768
             /gene="pro40"
             /number=1
     variation  3060
             /gene="pro40"
             /note="premature stop codon which causes defects in
             fruiting body maturation"
             /evidence=experimental
             /replace="t"
     intron  3769..3830
             /gene="pro40"
             /number=1
     exon    3831..4360
             /gene="pro40"
```

ORIGIN   /number=2

Query Match   66.7%; Score 20; DB 15; Length 6002;
Best Local Similarity 71.4%;   Pred. No. 4.3e+02;
Matches 20; Conservative 3; Mismatches 5; Indels 0; Gaps 0;

```
Qy    1 CAGGCUACCGUAGAGCAUCAUGAUCCU 28
          ||||:|||| |:|||||| ||:|||||
Db 2284 CAGGCTACGCCGAAAGCATCATGACCGT 2311
```

RESULT 9
BX571965_10
WPCOMMENT
Sequence split into 41 fragments   LOCUS BX571965 Accession BX571965

Fragment Name	Begin	End
BX571965_00	1	110000
BX571965_01	100001	210000
BX571965_02	200001	310000
BX571965_03	300001	410000
BX571965_04	400001	510000
BX571965_05	500001	610000
BX571965_06	600001	710000
BX571965_07	700001	810000
BX571965_08	800001	910000
BX571965_09	900001	1010000
BX571965_10	1000001	1110000
BX571965_11	1100001	1210000
BX571965_12	1200001	1310000
BX571965_13	1300001	1410000
BX571965_14	1400001	1510000
BX571965_15	1500001	1610000
BX571965_16	1600001	1710000
BX571965_17	1700001	1810000
BX571965_18	1800001	1910000
BX571965_19	1900001	2010000
BX571965_20	2000001	2110000
BX571965_21	2100001	2210000
BX571965_22	2200001	2310000
BX571965_23	2300001	2410000
BX571965_24	2400001	2510000
BX571965_25	2500001	2610000
BX571965_26	2600001	2710000
BX571965_27	2700001	2810000
BX571965_28	2800001	2910000
BX571965_29	2900001	3010000
BX571965_30	3000001	3110000
BX571965_31	3100001	3210000
BX571965_32	3200001	3310000
BX571965_33	3300001	3410000
BX571965_34	3400001	3510000
BX571965_35	3500001	3610000
BX571965_36	3600001	3710000
BX571965_37	3700001	3810000
BX571965_38	3800001	3910000
BX571965_39	3900001	4010000
BX571965_40	4000001	4074542

Continuation (11 of 41) of BX571965 from base 1000001 (BX571965 Burkholderia pseudomallei

Query Match   66.7%; Score 20; DB 1; Length 110000;
Best Local Similarity 75.0%;   Pred. No. 5.3e+02;
Matches 21; Conservative 2; Mismatches 5; Indels 0; Gaps 0;

```
Qy     3 GGCUACGCGGUAGAGCAUCAUGAUCCUGT 30
           ||| ||||||||| ||| |||||||||||
Db 53276 GGCTACGGCTACAGCTCGCTGATCCTGT 53303
```

RESULT 10
CP000010_23/c
WPCOMMENT
Sequence split into 36 fragments   LOCUS CP000010 Accession CP000010

```
Fragment Name    Begin        End
CP000010_00            1       110000
CP000010_01       100001       210000
CP000010_02       200001       310000
CP000010_03       300001       410000
CP000010_04       400001       510000
CP000010_05       500001       610000
CP000010_06       600001       710000
CP000010_07       700001       810000
CP000010_08       800001       910000
CP000010_09       900001      1010000
CP000010_10      1000001      1110000
CP000010_11      1100001      1210000
CP000010_12      1200001      1310000
CP000010_13      1300001      1410000
CP000010_14      1400001      1510000
CP000010_15      1500001      1610000
CP000010_16      1600001      1710000
CP000010_17      1700001      1810000
CP000010_18      1800001      1910000
CP000010_19      1900001      2010000
CP000010_20      2000001      2110000
CP000010_21      2100001      2210000
CP000010_22      2200001      2310000
CP000010_23      2300001      2410000
CP000010_24      2400001      2510000
CP000010_25      2500001      2610000
CP000010_26      2600001      2710000
CP000010_27      2700001      2810000
CP000010_28      2800001      2910000
CP000010_29      2900001      3010000
CP000010_30      3000001      3110000
CP000010_31      3100001      3210000
CP000010_32      3200001      3310000
CP000010_33      3300001      3410000
CP000010_34      3400001      3510000
CP000010_35      3500001      3510148
```

Continuation (24 of 36) of CP000010 from base 2300001 (CP000010 Burkholderia mallei ATCC

```
Query Match              66.7%;    Score 20;    DB 1;    Length 110000;
Best Local Similarity    75.0%;    Pred. No. 5.3e+02;
Matches   21;   Conservative   2;   Mismatches   5;   Indels   0;   Gaps   0;

Qy       3 GGCUACGCGUAGAGCAUCAUGAUCCUGU 30
           ||  ||| ||||||||:|||:||||||
Db   32065 GGCTACGCGTACAGTCGCTGATCCTGT 32038
```

RESULT 11
AC150398
LOCUS   AC150398   163878 bp   DNA   linear   INV 12-NOV-2004
DEFINITION Branchiostoma floridae clone CH302-18L14, complete sequence.
ACCESSION   AC150398
VERSION   AC150398.3 GI:55701323
KEYWORDS   HTG.
SOURCE   Branchiostoma floridae (Florida lancelet)
  ORGANISM Branchiostoma floridae
            Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
            Branchiostoma.
REFERENCE   1 (bases 1 to 163878)
  AUTHORS   DOE Joint Genome Institute and Stanford Human Genome Center.
  TITLE   Direct Submission
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 163878)
  AUTHORS   DOE Joint Genome Institute.
  TITLE   Direct Submission
  JOURNAL   Submitted (20-JUL-2004) Production Genomics Facility, DOE Joint
            Genome Institute, 2800 Mitchell Drive B100, Walnut Creek, CA
            94598-1698, USA
REFERENCE   3 (bases 1 to 163878)
  AUTHORS   Stanford Human Genome Center.
  CONSRTM   DOE Joint Genome Institute
  TITLE   Direct Submission

  JOURNAL   Submitted (28-OCT-2004) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
REFERENCE   4 (bases 1 to 163878)
  AUTHORS   Stanford Human Genome Center.
  CONSRTM   DOE Joint Genome Institute
  TITLE   Direct Submission
  JOURNAL   Submitted (12-NOV-2004) DOE Joint Genome Institute, 2800 Mitchell
            Drive, Walnut Creek, CA 94598, USA
COMMENT   On Nov 12, 2004 this sequence version replaced gi:54697207.
            Draft Sequence Produced by DOE Joint Genome Institute
            www.jgi.doe.gov
            Finishing Completed at Stanford Human Genome Center
            www-shgc.stanford.edu
            Quality: Phrap Quality >=40 100% of Sequence;
            Estimated Total Number of Errors is 0.
FEATURES   Location/Qualifiers
     source   1..163878
                     /organism="Branchiostoma floridae"
                     /mol_type="genomic DNA"
                     /db_xref="taxon:7739"
                     /clone="CH302-18L14"
ORIGIN

```
Query Match              66.7%;    Score 20;    DB 2;    Length 163878;
Best Local Similarity    71.4%;    Pred. No. 5.5e+02;
Matches   20;   Conservative   3;   Mismatches   5;   Indels   0;   Gaps   0;

Qy       3 GGCUACGCGUAGAGCAUCAUGAUCCUGT 30
           ||  :|| ||||||||:|| |||:||||
Db   68124 GGTTACTGGTAGAGCATCAGCATCCTGT 68151
```

RESULT 12
AC155516/c
LOCUS   AC155516   172391 bp   DNA   linear   HTG 25-JAN-2005
DEFINITION Zea mays strain B73 clone ZMMBBc0059P07, *** SEQUENCING IN PROGRESS
            ***, 22 unordered pieces.
ACCESSION   AC155516
VERSION   AC155516.2 GI:58082376
KEYWORDS   HTG; HTGS_PHASE1.
SOURCE   Zea mays
  ORGANISM Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 172391)
  AUTHORS   Chan,A.P., Pertea,G., Zheng,L., Cheung,F., Lee,D., Koo,H.,
            Utterback,T.R., Feldblyum,T.V., Rabinowicz,P., Fraser,C.M.,
            Schubert,K., SanMiguel,P., Ma,J., Pontaroli,A.C., Rohlfing,T.,
            Budiman,A., Bedell,J.A., Lakey,N., Barbazuk,B., Bennetzen,J. and
            Quackenbush,J.
  TITLE   Consortium for Maize Genomics - BAC skim sequencing and assembly
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 172391)
  AUTHORS   Chan,A.P., Pertea,G., Zheng,L., Cheung,F., Lee,D., Koo,H.,
            Utterback,T.R., Feldblyum,T.V., Rabinowicz,P., Fraser,C.M.,
            Schubert,K., SanMiguel,P., Ma,J., Pontaroli,A.C., Rohlfing,T.,
            Budiman,A., Bedell,J.A., Lakey,N., Barbazuk,B., Bennetzen,J. and
            Quackenbush,J.
  TITLE   Direct Submission
  JOURNAL   Submitted (15-JAN-2005) The Institute for Genomic Research (TIGR),
            9712 Medical Center Dr, Rockville, MD 20850
REFERENCE   3 (bases 1 to 172391)
  AUTHORS   Chan,A.P., Pertea,G., Zheng,L., Cheung,F., Lee,D., Koo,H.,
            Utterback,T.R., Feldblyum,T.V., Rabinowicz,P., Fraser,C.M.,
            Schubert,K., SanMiguel,P., Ma,J., Pontaroli,A.C., Rohlfing,T.,
            Budiman,A., Bedell,J.A., Lakey,N., Barbazuk,B., Bennetzen,J. and
            Quackenbush,J.
  TITLE   Direct Submission
  JOURNAL   Submitted (25-JAN-2005) The Institute for Genomic Research (TIGR),
            9712 Medical Center Dr, Rockville, MD 20850
COMMENT   On Jan 25, 2005 this sequence version replaced gi:57863037.
            ---------- Trace submission

```
Center name: TIGR
Seq_lib_id: ZGEV
--------- Project informaton
Web site: http://www.tigr.org/tdb/tgi/maize/
Contact: maize@tigr.org
----------
* NOTE: This is a 'working draft' sequence. It currently
* consists of 22 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*      1     2176: contig of 2176 bp in length
*   2177     2276: gap of unknown length
*   2277     8721: contig of 6445 bp in length
*   8722     8821: gap of unknown length
*   8822    10663: contig of 1842 bp in length
*  10664    10763: gap of unknown length
*  10764    24961: contig of 14198 bp in length
*  24962    25061: gap of unknown length
*  25062    27709: contig of 2648 bp in length
*  27710    27809: gap of unknown length
*  27810    38917: contig of 11108 bp in length
*  38918    39017: gap of unknown length
*  39018    40262: contig of 1245 bp in length
*  40263    40362: gap of unknown length
*  40363    42644: contig of 2282 bp in length
*  42645    42744: gap of unknown length
*  42745    63100: contig of 20356 bp in length
*  63101    63200: gap of unknown length
*  63201    67550: contig of 4350 bp in length
*  67551    67650: gap of unknown length
*  67651    69774: contig of 2124 bp in length
*  69775    69874: gap of unknown length
*  69875    71410: contig of 1536 bp in length
*  71411    71510: gap of unknown length
*  71511    74430: contig of 2920 bp in length
*  74431    74530: gap of unknown length
*  74531    79101: contig of 4571 bp in length
*  79102    79201: gap of unknown length
*  79202    80763: contig of 1562 bp in length
*  80764    80863: gap of unknown length
*  80864    83263: contig of 2400 bp in length
*  83264    83363: gap of unknown length
*  83364    89602: contig of 6239 bp in length
*  89603    89702: gap of unknown length
*  89703    97967: contig of 8265 bp in length
*  97968    98067: gap of unknown length
*  98068   112608: contig of 14541 bp in length
* 112609   112708: gap of unknown length
* 112709   152216: contig of 39508 bp in length
* 152217   152316: gap of unknown length
* 152317   160007: contig of 7691 bp in length
* 160008   160107: gap of unknown length
* 160108   172391: contig of 12284 bp in length.

FEATURES             Location/Qualifiers
     source          1..172391
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBBc0059P07"
     gap             2177..2276
                     /estimated_length=unknown
     gap             8722..8821
                     /estimated_length=unknown
     gap             10664..10763
                     /estimated_length=unknown
     gap             24962..25061
                     /estimated_length=unknown
     gap             27710..27809
                     /estimated_length=unknown
     gap             38918..39017
                     /estimated_length=unknown
     gap             40263..40362
                     /estimated_length=unknown
     gap             42645..42744
                     /estimated_length=unknown
     gap             63101..63200
                     /estimated_length=unknown
     gap             67551..67650
                     /estimated_length=unknown
     gap             69775..69874
                     /estimated_length=unknown
     gap             71411..71510
                     /estimated_length=unknown
     gap             74431..74530
                     /estimated_length=unknown
     gap             79102..79201
                     /estimated_length=unknown
     gap             80764..80863
                     /estimated_length=unknown
     gap             83264..83363
                     /estimated_length=unknown
     gap             89603..89702
                     /estimated_length=unknown
     gap             97968..98067
                     /estimated_length=unknown
     gap             112609..112708
                     /estimated_length=unknown
     gap             152217..152316
                     /estimated_length=unknown
     gap             160008..160107
                     /estimated_length=unknown
ORIGIN

Query Match     66.7%;  Score 20;  DB 14;  Length 172391;
Best Local Similarity  75.0%;  Pred. No. 5.5e+02;
Matches  21;  Conservative  2;  Mismatches  5;  Indels  0;  Gaps  0;

Qy       2 AGGCUACGGCTAGAGCAUCATGATCCUG 29
           || |||||| |||  ||||| |||||| :|
Db   96357 AGACAACGGCTCGAGCATTATGATGCTG 96330


RESULT 13
AC145730
LOCUS       AC145730    174912 bp    DNA    linear    HTG 16-JUN-2004
DEFINITION  Zea mays clone ZMMBBC0036A01, *** SEQUENCING IN PROGRESS ***, 2
            ordered pieces.
ACCESSION   AC145730
VERSION     AC145730.8  GI:48762548
KEYWORDS    HTG; HTGS_PHASE2; HTGS_FULLTOP; HTGS_ACTIVEFIN.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1 (bases 1 to 174912)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Butler,E., Wing,R., Bharti,A.K.
            and Messing,J.
  TITLE     Zea mays, Clone ZMMBBC0036A01
  JOURNAL   Unpublished
REFERENCE   2 (bases 1 to 174912)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Abouelleil,A., Allen,N.,
            Anderson,M., Arachchi,H.M., Barna,N., Bastien,V., Bloom,T.,
            Boguslavkiy,L., Boukhgalter,B., Camarata,J., Chang,J., Choepel,Y.,
            Collymore,A., Cook,A., Cooke,P., Corum,B., DeArellano,K.,
            Diaz,J.S., Dodge,S., Dooley,K., Dorris,L., Erickson,J., Faro,S.,
            Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S.,
            Graham,L., Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B.,
            Hall,J., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C.,
            Kamat,A., Karatas,A., Kells,C., Landers,T., Levine,R.,
            Lindblad-Toh,K., Liu,X., Lui,A., Mabbitt,R., MacLean,C.,
```

```
                Macdonald,P., Major,J., Manning,J., Matthews,C., McCarthy,M.,
                Meldrim,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J.,
                Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
                O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
                Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
                Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,
                Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
                Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Wu,X.,
                Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
                Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (23-JUL-2003) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
REFERENCE   3  (bases 1 to 174912)
  AUTHORS   Birren,B., Nusbaum,C., Lander,E., Butler,E., Wing,R., Bharti,A.K.,
            Messing,J., Abouelleil,A., Allen,N., Anderson,M., Anderson,S.,
            Arachchi,H.M., Barna,N., Bastien,V., Bloom,T., Boguslavkiy,L.,
            Boukhgalter,B., Camarata,J., Chang,Y.L., Choepel,Y., Collymore,A.,
            Cook,A., Cooke,P., Corum,B., DeArellano,K., Diaz,J.S., Dodge,S.,
            Dooley,K., Dorris,L., Erickson,J., Faro,S., Ferreira,P.,
            FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Graham,J.,
            Grand-Pierre,N., Hafez,N., Hagopian,D., Hagos,B., Hall,J.,
            Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A.,
            Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
            Liu,G., Liu,X., Lui,A., Mabbit,R., MacLean,C., Macdonald,P.,
            Major,J., Manning,J., Matthews,C., McCarthy,M., Meldrim,J.,
            Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C.,
            Nguyen,C., Nicol,R., Norbu,C., O'Connor,T., O'Donnell,P.,
            O'Neil,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N.,
            Rachupka,A., Ramasamy,U., Raymond,C., Retta,R., Rise,C., Rogov,P.,
            Roman,J., Schauer,S., Schupback,R., Seaman,S., Severy,P., Smith,C.,
            Spencer,B., Stange-Thomann,N., Stojanovic,N., Stubbs,M.,
            Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Wu,X.,
            Vassiliev,H., Venkataraman,V.S., Viel,R., Vo,A., Wilson,B., Wu,X.,
            Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
  TITLE     Direct Submission
  JOURNAL   Submitted (16-JUN-2004) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
  COMMENT   On Jun 16, 2004 this sequence version replaced gi:48717645.
            All repeats were identified using RepeatMasker.
            Smit, A.F.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RM/RepeatMasker.html
            ------------------- Genome Center
            Center: Whitehead Institute/MIT Center for Genome Research
            Center code: WIBR
            Web site: http://www-seq.wi.mit.edu
            Contact: sequence_submissions@broad.mit.edu
            -------------------
            Bharti, AK and Messing, J: The Plant Genome Initiative at
            Rutgers, Waksman Institute, Rutgers, The State University of New
            Jersey, 190 Frelinghuysen Road, Piscataway, NJ 08854, USA
            (http://pgir.rutgers.edu)
            Butler, E and Wing, R: Arizona Genomics Institute, Biological
            Sciences West, 448A, P.O. Box 210088, University of Arizona,
            Tucson, AZ 85721, USA (http://www.genome.arizona.edu)
            ------------------- Project Information
            Center project name: L29606
            Center clone name: 36_A1
            ------------------- Consensus Information
            This consensus is derived from a shotgun assembly that has been
            manually curated. It is the best representation of the BAC that we
            can generate without further laboratory work. The draft assembly
            has been edited, and if possible, ends identified by vector as well
            as by BAC end sequences, and contigs ordered and oriented. Bases
            that are not Ns are either above Q20 or manually edited. This
            assembly was performed with Arachne (Genome Res. 2002 12: 177-189).
            Genome Res. 2003 13: 91-96). All trace files for this project are
            available at the NCBI trace repository.
            (http://www.ncbi.nlm.nih.gov/Traces/trace.fcgi?). An exact list of
            reads used in this assembly are available at
            http://www.broad.mit.edu/annotation/plants/maize/randomclones.html.
            -------------------
          * NOTE: This is a 'working draft' sequence. It currently
```

```
              * consists of 2 contigs. Gaps between the contigs
              * are represented as runs of N. The order of the pieces
              * is believed to be correct as given, however the sizes
              * of the gaps between them are based on estimates that have
              * provided by the submittor.
              * This sequence will be replaced
              * by the finished sequence as soon as it is available and
              * the accession number will be preserved.
              *     1   28758: contig of 28758 bp in length
              * 28759   28858: gap of 100 bp
              * 28859  174912: contig of 146054 bp in length.
FEATURES             Location/Qualifiers
     source          1..174912
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /cultivar="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBBc0036A01"
                     /clone_lib="CHORI-201 Maize B73 BAC EcoRI Library"
     gap             28759..28858
                     /estimated_length=100
ORIGIN

Query Match             66.7%;  Score 20;  DB 14;  Length 174912;
Best Local Similarity   75.0%;  Pred. No. 5.5e+02;
Matches  21;  Conservative   2;  Mismatches   5;  Indels   0;  Gaps   0;

Qy        2 AGGCUACGGGTAGAGCAUCAUGATCCUG 29
            || || |||||| |||||  |||| | |:|
Db    25981 AGACAACGCGTCGAGCATTATGATGCTG 26008

RESULT 14
AC155574/c
LOCUS       AC155574              181826 bp    DNA     linear   HTG 25-JAN-2005
DEFINITION  Zea mays strain B73 clone ZMMBBc0177I11, *** SEQUENCING IN PROGRESS
            ***, 33 unordered pieces.
ACCESSION   AC155574
VERSION     AC155574.2  GI:58082433
KEYWORDS    HTG; HTGS_PHASE1.
SOURCE      Zea mays
  ORGANISM  Zea mays
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; PACCAD
            clade; Panicoideae; Andropogoneae; Zea.
REFERENCE   1  (bases 1 to 181826)
  AUTHORS   Chan,A.P., Pertea,G., Zheng,L., Cheung,F., Lee,D., Koo,H.,
            Utterback,T.R., Feldblyum,T.V., Rabinowicz,P., Fraser,C.M.,
            Schubert,K., SanMiguel,P., Ma,J., Pontaroli,A.C., Rohlfing,T., and
            Budiman,A., Bedell,J.A., Lakey,N., Barbazuk,B., Bennetzen,J. and
            Quackenbush,J.
  TITLE     Consortium for Maize Genomics - BAC skim sequencing and assembly
  JOURNAL   Unpublished
REFERENCE   2  (bases 1 to 181826)
  AUTHORS   Chan,A.P., Pertea,G., Zheng,L., Cheung,F., Lee,D., Koo,H.,
            Utterback,T.R., Feldblyum,T.V., Rabinowicz,P., Fraser,C.M.,
            Schubert,K., SanMiguel,P., Ma,J., Pontaroli,A.C., Rohlfing,T., and
            Budiman,A., Bedell,J.A., Lakey,N., Barbazuk,B., Bennetzen,J. and
            Quackenbush,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (15-JAN-2005) The Institute for Genomic Research (TIGR),
            9712 Medical Center Dr, Rockville, MD 20850
REFERENCE   3  (bases 1 to 181826)
  AUTHORS   Chan,A.P., Pertea,G., Zheng,L., Cheung,F., Lee,D., Koo,H.,
            Utterback,T.R., Feldblyum,T.V., Rabinowicz,P., Fraser,C.M.,
            Schubert,K., SanMiguel,P., Ma,J., Pontaroli,A.C., Rohlfing,T., and
            Budiman,A., Bedell,J.A., Lakey,N., Barbazuk,B., Bennetzen,J. and
            Quackenbush,J.
  TITLE     Direct Submission
  JOURNAL   Submitted (25-JAN-2005) The Institute for Genomic Research (TIGR),
            9712 Medical Center Dr, Rockville, MD 20850
  COMMENT   On Jan 25, 2005 this sequence version replaced gi:57863095.
            ------------------- Trace submission
```

```
Center name: TIGR
Seq_lib_id: ZGBP
------ Project informaton
Web site: http://www.tigr.org/tdb/tgi/maize/
Contact: maize@tigr.org
------
* NOTE: This is a 'working draft' sequence. It currently
* consists of 33 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*      1    2171: contig of 2171 bp in length
*   2172    2271: gap of unknown length
*   2272    3320: contig of 1049 bp in length
*   3321    3420: gap of unknown length
*   3421   14767: contig of 11347 bp in length
*  14768   14867: gap of unknown length
*  14868   15621: contig of 754 bp in length
*  15622   15721: gap of unknown length
*  15722   33188: contig of 17467 bp in length
*  33189   33288: gap of unknown length
*  33289   38383: contig of 5095 bp in length
*  38384   38483: gap of unknown length
*  38484   44947: contig of 6464 bp in length
*  44948   45047: gap of unknown length
*  45048   49262: contig of 4215 bp in length
*  49263   49362: gap of unknown length
*  49363   51474: contig of 2112 bp in length
*  51475   51574: gap of unknown length
*  51575   52995: contig of 1421 bp in length
*  52996   53095: gap of unknown length
*  53096   55245: contig of 2150 bp in length
*  55246   55345: gap of unknown length
*  55346   59760: contig of 4415 bp in length
*  59761   59860: gap of unknown length
*  59861   65081: contig of 5221 bp in length
*  65082   65181: gap of unknown length
*  65182   72066: contig of 6885 bp in length
*  72067   72166: gap of unknown length
*  72167   76345: contig of 4179 bp in length
*  76346   76445: gap of unknown length
*  76446   78368: contig of 1923 bp in length
*  78369   78468: gap of unknown length
*  78469   89004: contig of 10536 bp in length
*  89005   89104: gap of unknown length
*  89105   92932: contig of 3828 bp in length
*  92933   93032: gap of unknown length
*  93033  100121: contig of 7089 bp in length
* 100122  100221: gap of unknown length
* 100222  101693: contig of 1472 bp in length
* 101694  101793: gap of unknown length
* 101794  105200: contig of 3407 bp in length
* 105201  105300: gap of unknown length
* 105301  108453: contig of 3153 bp in length
* 108454  108553: gap of unknown length
* 108554  113771: contig of 5218 bp in length
* 113772  113871: gap of unknown length
* 113872  120608: contig of 6737 bp in length
* 120609  120708: gap of unknown length
* 120709  122835: contig of 2127 bp in length
* 122836  122935: gap of unknown length
* 122936  151409: contig of 28474 bp in length
* 151410  151509: gap of unknown length
* 151510  158009: contig of 6500 bp in length
* 158010  158109: gap of unknown length
* 158110  162836: contig of 4727 bp in length
* 162837  162936: gap of unknown length
* 162937  164885: contig of 1949 bp in length
* 164886  164985: gap of unknown length
* 164986  166737: contig of 1752 bp in length
* 166738  166837: gap of unknown length
* 166838  168126: contig of 1289 bp in length
* 168127  168226: gap of unknown length
* 168227  170510: contig of 2284 bp in length
* 170511  170610: gap of unknown length
* 170611  181826: contig of 11216 bp in length.

FEATURES             Location/Qualifiers
     source          1..181826
                     /organism="Zea mays"
                     /mol_type="genomic DNA"
                     /strain="B73"
                     /db_xref="taxon:4577"
                     /clone="ZMMBBc0177I11"
     gap             2172..2271
                     /estimated_length=unknown
     gap             3321..3420
                     /estimated_length=unknown
     gap             14768..14867
                     /estimated_length=unknown
     gap             15622..15721
                     /estimated_length=unknown
     gap             33189..33288
                     /estimated_length=unknown
     gap             38384..38483
                     /estimated_length=unknown
     gap             44948..45047
                     /estimated_length=unknown
     gap             49263..49362
                     /estimated_length=unknown
     gap             51475..51574
                     /estimated_length=unknown
     gap             52996..53095
                     /estimated_length=unknown
     gap             55246..55345
                     /estimated_length=unknown
     gap             59761..59860
                     /estimated_length=unknown
     gap             65082..65181
                     /estimated_length=unknown
     gap             72067..72166
                     /estimated_length=unknown
     gap             76346..76445
                     /estimated_length=unknown
     gap             78369..78468
                     /estimated_length=unknown
     gap             89005..89104
                     /estimated_length=unknown
     gap             92933..93032
                     /estimated_length=unknown
     gap             100122..100221
                     /estimated_length=unknown
     gap             101694..101793
                     /estimated_length=unknown
     gap             105201..105300
                     /estimated_length=unknown
     gap             108454..108553
                     /estimated_length=unknown
     gap             113772..113871
                     /estimated_length=unknown
     gap             120609..120708
                     /estimated_length=unknown
     gap             122836..122935
                     /estimated_length=unknown
     gap             151410..151509
                     /estimated_length=unknown
     gap             158010..158109
                     /estimated_length=unknown
     gap             162837..162936
                     /estimated_length=unknown
     gap             164886..164985
                     /estimated_length=unknown
     gap             166738..166837
                     /estimated_length=unknown
```

```
gap              168127..168226
                 /estimated_length=unknown
gap              170511..170610
                 /estimated_length=unknown
ORIGIN

Query Match      66.7%;  Score 20;  DB 14;  Length 181826;
Best Local Similarity  75.0%;  Pred. No. 5.5e+02;
Matches  21;  Conservative  2;  Mismatches  5;  Indels  0;  Gaps  0;

Qy       2 AGGCUACGCGUAGAGCAUCAUCAUGAUCCUG 29
           || ||||||| |||||| ||||| |||
Db   26450 AGACAAACGCGTCGAGCATTATGATGCTG 26423

RESULT 15
AE004320
LOCUS       AE004320          10205 bp    DNA    linear   BCT 10-OCT-2003
DEFINITION  Vibrio cholerae O1 biovar eltor str. N16961 chromosome I, section
            228 of 251 of the complete chromosome.
ACCESSION   AE004320 AE003852
VERSION     AE004320.1  GI:9657082
KEYWORDS    .
SOURCE      Vibrio cholerae O1 biovar eltor str. N16961
  ORGANISM  Vibrio cholerae O1 biovar eltor str. N16961
            Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
            Vibrionaceae; Vibrio.
REFERENCE   1  (bases 1 to 10205)
  AUTHORS   Heidelberg,J.F., Eisen,J.A., Nelson,W.C., Clayton,R.A., Gwinn,M.L.,
            Dodson,R.J., Haft,D.H., Hickey,E.K., Peterson,J.D., Umayam,L.,
            Gill,S.R., Nelson,K.E., Read,T.D., Tettelin,H., Richardson,D.,
            Ermolaeva,M.D., Vamathevan,J., Bass,S., Qin,H., Dragoi,I.,
            Sellers,P., McDonald,L., Utterback,T., Fleishmann,R.D.,
            Nierman,W.C. and White,O.
  TITLE     DNA sequence of both chromosomes of the cholera pathogen Vibrio
            cholerae
  JOURNAL   Nature 406 (6795), 477-483 (2000)
   PUBMED   10952301
REFERENCE   2  (bases 1 to 10205)
  AUTHORS   Heidelberg,J.F., Eisen,J.A., Nelson,W.C., Clayton,R.A., Gwinn,M.L.,
            Dodson,R.J., Haft,D.H., Hickey,E.K., Peterson,J.D., Umayam,L.A.,
            Gill,S.R., Nelson,K.E., Read,T.D., Tettelin,H., Richardson,D.,
            Ermolaeva,M.D., Vamathevan,J., Bass,S., Qin,H., Dragoi,I.,
            Sellers,P., McDonald,L., Utterback,T., Fleishmann,R.D.,
            Nierman,W.C., White,O., Salzberg,S.L., Smith,H.O., Colwell,R.R.,
            Mekalanos,J.J., Venter,J.C. and Fraser,C.M.
  TITLE     Direct Submission
  JOURNAL   Submitted (14-JUN-2000) The Institute for Genomic Research, 9712
            Medical Center Dr, Rockville, MD 20850, USA
FEATURES             Location/Qualifiers
     source          1..10205
                     /organism="Vibrio cholerae O1 biovar eltor str. N16961"
                     /mol_type="genomic DNA"
                     /strain="N16961"
                     /serotype="O1"
                     /db_xref="taxon:243277"
                     /chromosome="I"
                     /note="biotype: El Tor"
     gene            128..1639
                     /gene="VC2501"
     CDS             128..1639
                     /gene="VC2501"
                     /note="similar to SP:P11648 GB:X15130 PID:1054725
                     PID:43309 PID:537102; identified by sequence similarity;
                     putative"
                     /codon_start=1
                     /transl_table=11
                     /product="aminopeptidase A/I"
                     /protein_id="AAF95643.1"
                     /db_xref="GI:9657083"
                     /translation="MEFSVKSGSPEKQRSACIVVGVFEPRRLSPVAEQLDKISDGYIS
                     SLLRRGDLEGKPGQMLLLHQVPGVLSERVLLVGCGKERELGERQYKEIIQKTINTLNE
                     TGGSMEAVCFLTELHVKGRDTYWKVRQAVEATKDGLYIFDQFKSVKPEIRRPLRKLVFN
```

```
gene            1723..2172
                /gene="VC2502"
CDS             1723..2172
                /gene="VC2502"
                /note="similar to SP:P28905 SP:P11649 GB:L04574 GB:X15130
                GB:Z14155; identified by sequence similarity; putative"
                /codon_start=1
                /transl_table=11
                /product="DNA polymerase III, chi subunit"
                /protein_id="AAF95644.1"
                /db_xref="GI:9657084"
                /translation="MPTATFYLISRDSAQATPEGLAQYVLVLTRHFAAQGARIYLQCQ
                DKPHAEQLAERFWQIEPEQFLAHNLVEEGPKSGTAIEIGYVGVSPSWNRQLAINLADN
                QTTFARTFTEVVDFVPCAENAKQLARERYKLYRQAGYQLQTVEIQYP"

gene            2266..5127
                /gene="VC2503"
CDS             2266..5127
                /gene="VC2503"
                /note="similar to SP:P07118 GB:J03497 GB:X05891 PID:147906
                PID:43302; identified by sequence similarity; putative"
                /codon_start=1
                /transl_table=11
                /product="valyl-tRNA synthetase"
                /protein_id="AAF95645.1"
                /db_xref="GI:9657085"
                /translation="MEKTYNPTSIEQDLYKTWEEQGYFKPHGDTSKDAYSIMIPPPNV
                TGSLHMGHAFODTIMDTLIRCQRMKGKNTLWQVGTDHAGIATOMVERKIAAEEGKTK
                HDYGRDAFIDKIWEWKAESGGTITKQLRRLGASVDWDRERFTMDDGFYKAVQEVFVRL
                YKDDLIYRGKRLVNWDPKLHTAISDLEVENKETKGHMWHFRYPLADGVKTADGKDYIV
                VATTRPETMLGDTGAVNPEDPRYKDLIGKEIILPIVGRRIPIVGDEHADMEKGTGCYV
                KITPAHDFNDYEVGKRHQLPMINILTFDANIRDAAEVFNSNGEASNAYGTEIPAKYQG
                MERFAARKAIVAEFEELGLLQEIKDHDLTVPYGDRGGVVIEPMLTDQWYVRAGILAKP
                AVEAVENGDIQFPKQYENMYFSWMRDIQDWCISRQLWWGHRIPAWYDBQGNVFVGRN
                EEEVRAENNIAADVALRQDDDVLDTWFSSALWTFGTLGWPEKTPELKVFHPTDVLVTG
                FDIIFFWVARMIMMTMHFCKDEDGKAQVPFKTVYVTGLIRDENGDKMSKSKGNVLDPI
                DMIDGIDLESLVAKRTGNMMOPQLAAKIEKNTRKTFENGIEAYGTDSLRFTLAAMAST
                GRDINWDMKRLEGYRNFCNKLWNASRYVLMNTEEQDCGFAAGAELEYSLADKWIESQF
                ELAAKEPNGHIDNFRLDMAANTLYEFTWNQFCDWYLELTKPVLWKGTEAQQRATRRTL
                ITVLEKTLRLAHPVIPYITERIWQSVKPLVDGVEGDTIMLQALPQYDVVANFNQEALDD
                IEWVKAFITSIRNLRAEYDINPGKPLEVMLKAANEQDAARIEANKPVLVSLAKLESIR
                VLADGEATPACATALVGKSELMIPMAGLIDKDAELDRLAKEIAKTQGEIARIEGKLGN
                EGFVARAPEAVITKERELKLAGYQEALVKLEQQKATIRAL"

gene            complement(5205..6182)
                /gene="VC2504"
CDS             complement(5205..6182)
                /gene="VC2504"
                /note="similar to GB:L42023 SP:P45250 PID:1007767
                PID:1221703 PID:1205791; identified by sequence
                similarity; putative"
                /codon_start=1
                /transl_table=11
                /product="2-hydroxyacid dehydrogenase family protein"
                /protein_id="AAF95646.1"
                /db_xref="GI:9657086"
                /translation="MSLPTTSLPTVVFLDRATIPRHISLPALPFEHHWLEYDACEPQQ
                VVERLLAADIVITNKVVLTREMLIQLPKLKLIAISATGTNNVDLPACRDLNIAVCNVQ
                GYATRSVPEHVVAMMFALRRNLIGYHNDLAAGEWQRHKQFCEFTHPIGDIAGSTMGII
                GSGALGQATANLARALGMHVLLAERKGQVECRDGYTSFEQVLAQSDVLSLHCPLTDET
                RNIISEBELAQMNPNALLINTGRGGLVDEQALVDALKRRQIAGAGVDVFSAEPADMDN
                PLIANRDLPNLLLTPHVAWGSDSSIQQLATILIDNISAFMRGEAKNRVV"

gene            complement(6222..6959)
                /gene="VC2505"
CDS             complement(6222..6959)
                /gene="VC2505"
                /note="similar to SP:P39219 GB:U00096 PID:1786244;
                identified by sequence similarity; putative"
                /codon_start=1
```

```
VPTRRELNLGERAITHGLAISSGVKACKDLGNMPNIANPAYLASQARRLADDYESIT
TKIIGEEMEKLGMASYLAVGRGSRNESMMSVIEYKGNPDPEAKPIVLVGKGLTFDSG
GISLKPGEMDEMKYDMCGAASVFGTMKAIAKLGLPNVIGVLAGCENMPGSNAYRPG
DILTTMSGQTVEVLNTDAEGRIVLCDVLTYVERFEPECVVDVATLTGACVIALGHHIS
AVMSNINPLAHELVNASEQSSDRAWRLPLADEYHEQLKSPFADMANIGGRPGGAITAA
CFLSKFAKKYNWAHLDIAGTAWKSGAAKGSTGRPVSLLVQFLLNRSGGLDAEE"
```

```
              /transl_table=11
              /product="ribosomal large subunit pseudouridine synthase
              A"
              /protein_id="AAF95647.1"
              /db_xref="GI:9657087"
              /translation="MAMTEYNPPKDPWIDVIYEDEHIIAANKPSGLLSVPGRPVEHYD
              SMWARLVDYCPDVKVVHRLDMSTSGLILFAKGKHMESALKKQFQYRLTHKIYYARVWG
              VMEQNEGEVNLPLCCDWPNRPKQKVCFEEGRDSLIRYQVVKREEKTVVRLIPITGRS
              HQLRVHMQSLGHPIVGDDLYAPPEAVEFADRLQLHAAELCFYHPRSHWLRTLFVPCDF
              YPEAAEEQILQHFDPERKLPDYKKLSQD"
    gene     complement(7056..9965)
              /gene="VC2506"
    CDS      complement(7056..9965)
              /gene="VC2506"
              /note="similar to GB:U00096 PID:1786245; identified by
              sequence similarity; putative"
              /codon_start=1
              /transl_table=11
              /product="RNA polymerase-associated protein HepA"
              /protein_id="AAF95648.1"
              /db_xref="GI:9657088"
              /translation="MSFALGQRWISDTESDLGLGTVVALDARTVTLMFAASEENRVYA
              RSDAPVTRVIFNVGDVVDSQQGWSLQVEQVVEBDQGVYTYLGTRVDTBESGVALREIFL
              SNQIRFNKPQDKLFAGQIDRMDNFVLRYRALTNQYQQHKSPMRGLCGMRAGLIPHQLY
              IAHEVGRRHAPRVLLADEVGLGKTIEAGMIIHQQVLTGRAERILIVVPETLQHQMLVE
              MMRRFNLHFSIFDEERCVEAFSEADNPFETQQYVLCSLDFLRKSRQRFEQALEAEWDL
              LVVDEAHHLEWHPEKPSREYQVIEALAEQTPGVLLLTATPEQLGRESHFARLRLLDAD
              RFYDYEAFVKEEQYAPVADAVTALFSGEKLSDEAKNKITELLSEQDVEPLFKALESH
              ASEDEIALARQELIDNLMDRHGTGRVLFRNTRAAIKGFPVRNVHLLPLEIPSQYTTSM
              RVAGMLGGKLTPEARAMKMLYPEEIFQEFEGEBSSWWQFDSRVNWLLEKVKAKRSEKI
              LVIASRASTALQEQALREREGIRATVFHEGMSIIERDKAAAYFAQEEGAQVLICSE
              IGSEGRNFQFANQLVMFDLPFNPDLLEQRIGRLDRIGQKRDIDVVVPYLTETSQAILA
              RWFQEGLNAFAETCPTGRAVDAFAERLIPILAAGGEBLEVIIEESAKLNKTLKSQL
              EVGRDRLLEMHSNGEKAQQIAEQIAKTDGDTNLVTFALSLFDAIGLHQEDRGENALV
              VTPAEHMMVPSYPGLPYEGATITFDRDTALSREDMHFISWEHPMVQGGIDLLMSEGVG
              TCAVSLLKNKALPVGTILLELVYVVDAQAPKRSGISRFLPVSPIRILMDARGNDLSSQ
              VEFESFNRQLSPVNRHLASKLVSSVQHDVHRLITASETAVEPRVSAIREQAQRDMQQS
              LNSELERLLALKAVNPNIRDEEIEVLDQQIKELTGYIAQAQYQLDSLRLIVVAHN"

    ORIGIN

    Query Match      66.0%;   Score 19.8;  DB 1;   Length 10205;
    Best Local Similarity   87.0%;  Pred. No. 5.5e+02;
    Matches   20;  Conservative    1;  Mismatches    2;   Indels    0;  Gaps    0;

    Qy      7 ACGCGTAGAGCAUCATGATCCUG 29
              ||||||| ||||:|||||||||||
    Db   2363 ACGCGTACAGCATCATGATCCCG 2385

    Search completed: January  8, 2006, 16:24:45
    Job time : 1732 secs
```

THIS PAGE BLANK (USPTO)

```
            GenCore version 5.1.6
      Copyright (c) 1993 - 2006 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on:   January 8, 2006, 15:08:49  ; Search time 2518 Seconds
                                       (without alignments)
                                      557.431 Million cell updates/sec

Title:          US-10-791-367-1
Perfect score:  30
Sequence:       1 cagcuacgcgtagagcaucatgatccugt 30

Scoring table:  IDENTITY_NUC
                Gapop 10.0 , Gapext 1.0

Searched:       41078325 seqs, 2339541228 residues

Total number of hits satisfying chosen parameters:    82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
                 Maximum Match 100%
                 Listing first 45 summaries

Database :    EST:*
              1:  gb_est1:*
              2:  gb_est2:*
              3:  gb_est3:*
              4:  gb_htc:*
              5:  gb_est4:*
              6:  gb_est5:*
              7:  gb_est6:*
              8:  gb_est7:*
              9:  gb_gss1:*
             10:  gb_gss2:*
             11:  gb_gss3:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

                              SUMMARIES
```

```
Result                Query
No.    Score  Match  Length  DB  ID        Description
-----  -----  -----  ------  --  --------  -----------
c   1  20.6   68.7     389   5   BW778170  BW778170
    2  20.6   68.7     526   7   CO786247  BL285D_D0
    3  20.6   68.7     683   1   AU170138  AU170138
c   4  20.6   68.7     724   3   BJ716634  BJ716634
    5  20.4   68.0     451   9   BX306829  BX306829
c   6  20.4   68.0     501   9   CE251821  tigr-gss-
c   7  20.4   68.0     521   9   AQ405865  HS_5043_B
c   8  20.4   68.0     631   1   AJ745356  AJ745356
c   9  20.4   68.0     639   5   CA058120  ssalrgga51
c  10  20.4   68.0     759   9   CC494201  CH240_329
   11  20.4   68.0    1332   8   DN711869  CLJ97-H05
   12  20.4   68.0    1351   8   DN686483  CGX47-C07
c  13  20.4   68.0    1372   8   DN711868  CLJ97-H05
   14  20.4   68.0    1413   8   DN686482  CGX47-C07
c  15  20.4   68.0    1669  10   CL640951  CH213-1N0
c  16  20.2   67.3     362  10   C2914756  4013008E0
   17  20     66.7     582   9   DN407618  LIB4006-0
   18  20     66.7     697   9   CC145434  ZMMBBb000
   19  20     66.7     820   5   BU745550  CH1#010_E
c  20  20     66.7     862   9   CC982446  ZUAEB75TH
c  21  20     66.7     916   5   BU745549  CH1#010_E
c  22  19.6   65.3     393   5   BX565471  BX565471
```

```
c  23  19.6   65.3   500   5  BX558657  BX558657
c  24  19.6   65.3   520   7  CO294995  EK200126.
c  25  19.6   65.3   573   5  BX562437  BX562437
c  26  19.6   65.3   581   5  BX556173  BX556173
c  27  19.6   65.3   591   5  BX568347  BX568347
c  28  19.6   65.3   597   5  BX568861  BX568861
   29  19.6   65.3   637   5  BX551247  BX551247
c  30  19.6   65.3   837   1  AI404024  GH23607.5
   31  19.4   65.3  1153  10  AG385831  Mus muscu
c  32  19.4   64.7   414   9  AQ181478  HS_3219_A
   33  19.4   64.7   713   9  BH035476  RPCI-24-2
c  34  19.4   64.7   768   3  CC931167  ZMMBBc054
c  35  19.4   64.7   777   3  BJ566348  BJ56648
c  36  19.4   64.7   792   5  BU363359  60790407
c  37  19.4   64.7   810   9  AQ875939  V131C6 mT
c  38  19.4   64.7   840   9  AQ251296  T30I10-Sp
c  39  19.4   64.7  5463   4  CR858601  Pongo pyg
c  40  19.2   64.0   466   2  BE125931  DG1_59_G0
c  41  19.2   64.0   585   8  DR063891  ip7Ie04.g
   42  19.2   64.0   654   9  BF528338  60204251
c  43  19.2   64.0   666   9  BZ335768  h26g05.g
c  44  19.2   64.0   724  10  CW487118  fsbb001f2
   45  19.2   64.0   772   9  B2628162  ih58c10.b
```

ALIGNMENTS

```
RESULT 1
BW778170/c
LOCUS       BW778170          389 bp    mRNA     linear    EST 10-AUG-2005
DEFINITION  BW778170 Amphioxus Branchiostoma floridae unpublished cDNA library,
            gastrula whole animal Branchiostoma floridae cDNA clone bbga048o03
            5', mRNA sequence.
ACCESSION   BW778170
VERSION     BW778170.1  GI:6630664
KEYWORDS    EST.
SOURCE      Branchiostoma floridae (Florida lancelet)
  ORGANISM  Branchiostoma floridae
            Eukaryota; Metazoa; Chordata; Cephalochordata; Branchiostomidae;
            Branchiostoma.
REFERENCE   1  (bases 1 to 389)
  AUTHORS   Yu,J., Holland,L.Z., Shin-i,T., Kohara,Y., Satou,Y. and Satoh,N.
  TITLE     Expressed genes in Branchiostoma floridae
  JOURNAL   Unpublished (2005)
COMMENT     Contact: Tadasu Shin-i
            Center For Genetic Resource Information
            National Institute of Genetics
            1111 Yata, Mishima, Shizuoka 411-8540, Japan
            Tel: 81-559-81-6856
            Fax: 81-559-81-6855
            Email: tshini@genes.nig.ac.jp
            If you want to have a cDNA clone for this EST or if you have any
            questions, please send an e-mail to Nori Satoh
            (satoh@ascidian.zool.kyoto-u.ac.jp) and its cc to Yutaka Satou
            (yutaka@ascidian.zool.kyoto-u.ac.jp).
FEATURES             Location/Qualifiers
     source          1..389
                     /organism="Branchiostoma floridae"
                     /mol_type="mRNA"
                     /db_xref="taxon:7739"
                     /clone="bbga048o03"
                     /tissue_type="whole animal"
                     /dev_stage="gastrula"
                     /clone_lib="Amphioxus Branchiostoma floridae unpublished
                     cDNA library, gastrula whole animal"
ORIGIN

Query Match             68.7%;  Score 20.6;  DB 5;  Length 389;
Best Local Similarity   77.8%;  Pred. No. 2.5e+02;
Matches   21;  Conservative   2;  Mismatches    4;  Indels   0;  Gaps   0;

Qy         2 AGGCUACGCGCTAGAGCAUCATGATCCU 28
```

```
Db          ||||||:|| ||||||||| ||||
        195 AGGCTACACGTAGAGCAACTTGATGCT 169
```

RESULT 2
CO786247
LOCUS   CO786247   526 bp   mRNA   linear   EST 05-AUG-2004
DEFINITION   BL285D_D03 6-Day Axolotl Tail Blastema (6DAxBL) Ambystoma mexicanum
             cDNA 5' similar to Unknown (protein for MGC:89099) (Xenopus
             tropicalis), mRNA sequence.
ACCESSION   CO786247
VERSION   CO786247.1 GI:51002227
KEYWORDS   EST.
SOURCE   Ambystoma mexicanum (axolotl)
  ORGANISM   Ambystoma mexicanum
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Amphibia; Batrachia; Caudata; Salamandroidea; Ambystomatidae;
             Ambystoma.
REFERENCE   1 (bases 1 to 526)
  AUTHORS   Habermann,B., Bebin,A.G., Herklotz,S., Volkmer,M., Eckelt,K.,
             Pehlke,K., Epperlein,H.H., Schackert,H.K., Wiebe,G. and Tanaka,E.M.
  TITLE   An Ambystoma mexicanum EST sequencing project: Analysis of 17,352
             expressed sequence tags from embryonic and regenerating blastema
             cDNA libraries
  JOURNAL   Genome Biol. (2004) In press
COMMENT   Contact: Elly M. Tanaka
             Tanaka Lab
             Max Planck Institute of Molecular Cell Biology and Genetics,
             Dresden
             Pfotenhauerstrasse 108, 01307 Dresden, Germany
             Tel: 0049 351 210 2620
             Fax: 0049 351 210 1489
             Email: tanaka@mpi-cbg.de
             Plate: BL285D row: 03 column: D
             Seq primer: GCA CAT TAG GCC TAT TTA GGT GAC A.
FEATURES   Location/Qualifiers
     source 1..526
             /organism="Ambystoma mexicanum"
             /mol_type="mRNA"
             /db_xref="taxon:8296"
             /tissue_type="Tail Blastema"
             /cell_type="regenerating tail blastema"
             /clone_lib="6-Day Axolotl Tail Blastema (6DAxBL)"
             /note="Vector: pCMVSport6; Site 1: NotI; Site 2: SalI;
             Unnormalized cDNA plasmid library prepared by Invitrogen.
             Size fractionated mRNA was polydT primed and cloned into
             NotI-SalI site of pCMVsport6. Bacterial host is
             EMDH10B-TONA. Average insert size is 1.67 kB.
             TAG_LIB=6DAxBL"
ORIGIN

Query Match   68.7%; Score 20.6; DB 7; Length 526;
  Best Local Similarity   77.8%; Pred. No. 2.6e+02;
  Matches   21; Conservative   2; Mismatches   4; Indels   0; Gaps   0;

Qy   4 GCUACGCGUAGAGCAUCAUGAUCCUGT 30
         ||||||| |||||||:| ||||||||||
Db   230 GCTACACGGAAGAGCGGCATGATCCTGT 256

RESULT 3
AU170138
LOCUS   AU170138   683 bp   mRNA   linear   EST 24-FEB-2005
DEFINITION   AU170138 Ol-br-ad cDNA Oryzias latipes cDNA clone br5670, mRNA
             sequence.
ACCESSION   AU170138
VERSION   AU170138.1 GI:12592209
KEYWORDS   EST.
SOURCE   Oryzias latipes (Japanese medaka)
  ORGANISM   Oryzias latipes
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
             Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
             Beloniformes; Adrianichthyidae; Oryziinae; Oryzias; Oryzias.
REFERENCE   1 (bases 1 to 683)
  AUTHORS   Mita,K., Ishikawa,Y. and Yamauchi,M.
  TITLE   Establishment of cDNA database of medaka, Oryzias latipes
  JOURNAL   Unpublished (2001)
COMMENT   Contact: Mita K
             Genome Research Group
             National Institute of Agrobiological Sciences
             Owashi 1-2, Tsukuba, Ibaraki 305-8634, Japan
             Tel: 81-29-838-6120
             Fax: 81-29-838-6121
             Email: kmita@nias.affrc.go.jp
             method:uni-directional sequence direction:sequenced from T3 primer
             (5' -> 3').
FEATURES   Location/Qualifiers
     source 1..683
             /organism="Oryzias latipes"
             /mol_type="mRNA"
             /strain="HNI"
             /db_xref="taxon:8090"
             /clone="br5670"
             /sex="female/male mixed"
             /tissue_type="brain"
             /dev_stage="adult"
             /clone_lib="Ol-br-ad cDNA"
ORIGIN

Query Match   68.7%; Score 20.6; DB 1; Length 683;
  Best Local Similarity   74.1%; Pred. No. 2.7e+02;
  Matches   20; Conservative   3; Mismatches   4; Indels   0; Gaps   0;

Qy   4 GCUACGCGUAGAGCAUCAUGAUCCUGT 30
         |||||| ||||||||||||||||:||||
Db   483 GCTACAAGGAGAGCATCATGAGCCTGT 509

RESULT 4
BJ716634/c
LOCUS   BJ716634   724 bp   mRNA   linear   EST 08-MAR-2004
DEFINITION   BJ716634 MF01FFA cDNA Oryzias latipes cDNA clone MF01FFA013h19 3',
             mRNA sequence.
ACCESSION   BJ716634
VERSION   BJ716634.1 GI:4257475
KEYWORDS   EST.
SOURCE   Oryzias latipes (Japanese medaka)
  ORGANISM   Oryzias latipes
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
             Acanthomorpha; Acanthopterygii; Percomorpha; Atherinomorpha;
             Beloniformes; Adrianichthyidae; Oryziinae; Oryzias.
REFERENCE   1 (bases 1 to 724)
  AUTHORS   Kohara,Y., Shin-i,T., Kimura,T., Narita,T., Jindo,T. and Takeda,H.
  TITLE   Medaka EST Project in Takeda's lab
  JOURNAL   Unpublished (2001)
COMMENT   Contact: Tadasu Shin-i
             Center For Genetic Resource Information
             National Institute of Genetics
             1111 Yata, Mishima, Shizuoka 411-8540, Japan
             Tel: 81-559-81-6856
             Fax: 81-559-81-6855
             Email: tshini@genes.nig.ac.jp.
FEATURES   Location/Qualifiers
     source 1..724
             /organism="Oryzias latipes"
             /mol_type="mRNA"
             /strain="Hd-rR"
             /db_xref="taxon:8090"
             /clone="MF01FFA013h19"
             /sex="mixture of female and male"
             /tissue_type="whole embryo"
             /dev_stage="fry stage 40"
             /clone_lib="MF01FFA cDNA"
ORIGIN

```
Query Match       68.7%;   Score 20.6;  DB 3;  Length 724;
Best Local Similarity   74.4%;   Pred. No. 2.7e+02;
Matches   20;  Conservative   3;  Mismatches   4;  Indels   0;  Gaps   0;

Qy      4  GCUACGCGTAGAGCAUCATGATCCUGT 30
           ||:||  |  |||||||||||:||| ||:||
Db    605  GCTACAAGGAGAGCATCATGAGCCTGT 579
```

RESULT 5
BX306829

```
LOCUS       BX306829                451 bp    mRNA     linear   EST 11-MAY-2004
DEFINITION  BX306829 tcay Oncorhynchus mykiss cDNA clone tcay0017b.e.20 3prim,
            mRNA sequence.
ACCESSION   BX306829
VERSION     BX306829.2  GI:40223548
KEYWORDS    EST.
SOURCE      Oncorhynchus mykiss (rainbow trout)
ORGANISM    Oncorhynchus mykiss
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei;
            Protacanthopterygii; Salmoniformes; Salmonidae; Oncorhynchus.
REFERENCE   1 (bases 1 to 451)
AUTHORS     Govoroun,M., Guiguen,Y. and Le Gac,F.
TITLE       Construction and primary characterization of normalized cDNA
            libraries in rainbow trout, Oncorhynchus mykiss
JOURNAL     Unpublished (2003)
COMMENT     On Apr 7, 2003 this sequence version replaced gi:29587474.
            Contact: Guiguen Y
            INRA - SCRIBE
            Campus de beaulieu, RENNES cedex, 35042, France
            Tel: 02.23.48.50.09
            Fax: 02.23.48.50.20
            Email: Yann.Guiguen@beaulieu.rennes.inra.fr
            Sequence cleaned of vector, adaptator and repetitions. Contact us
            at sigenasupport@jouy.inra.fr to obtain the chromatogram of this
            sequence.
            Plate: 0017  row: e  column: 20
            Seq primer: M13F.
FEATURES    Location/Qualifiers
     source          1..451
                     /organism="Oncorhynchus mykiss"
                     /mol_type="mRNA"
                     /db_xref="taxon:8022"
                     /clone="tcay0017b.e.20"
                     /tissue_type="adipose tissue, blood, brain,
                     differentiating gonads, gills, interrenal, intestine,
                     kidney, liver, muscle, ovary, pituitary, testis"
                     /dev_stage="from embryos to adults"
                     /lab_host="DH10B"
                     /clone_lib="tcay"
                     /note="Vector: pT7T3D-pac; Rainbow trout multi-tissues -
                     normalized + 1 subtraction (tcay) ; Clone distribution :
                     AGENAE Resource centre. Francois PIUMI,
                     Francois.Piumi@jouy.inra.fr, INRA, CEA Radiobiologie et
                     Etude du genome (LREG), Domaine de Vilvert, 78352,
                     Jouy-en-Josas cedex, France, +33 (0) 1.34.65.28.02, +33
                     (0) 1.34.65.22.73"
ORIGIN

Query Match       68.0%;   Score 20.4;  DB 5;  Length 451;
Best Local Similarity   70.0%;   Pred. No. 3.1e+02;
Matches   21;  Conservative   3;  Mismatches   6;  Indels   0;  Gaps   0;

Qy      1  CAGGCUACGCGTAGAGCAUCATGATCCUGT 30
           ||||||  |  |||||||||:|||||  |:||
Db     79  CTGGCTACAAGAAGAGCATCATGGCCCTGT 108
```

RESULT 6
CE251821/c

```
LOCUS       CE251821                501 bp    DNA      linear   GSS 26-SEP-2003
DEFINITION  tigr-gss-dog-170003591549 Dog Library Canis familiaris genomic,
            genomic survey sequence.
ACCESSION   CE251821
VERSION     CE251821.1  GI:35948893
KEYWORDS    GSS.
SOURCE      Canis familiaris (dog)
ORGANISM    Canis familiaris
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Laurasiatheria; Carnivora; Fissipedia; Canidae;
            Canis.
REFERENCE   1 (bases 1 to 501)
AUTHORS     Kirkness,E.F., Bafna,V., Halpern,A.L., Levy,S., Remington,K.,
            Rusch,D.B., Delcher,A.L., Pop,M., Wang,W., Fraser,C.M. and
            Venter,J.C.
TITLE       The dog genome: survey sequencing and comparative analysis
JOURNAL     Science 301 (5641), 1898-1903 (2003)
PUBMED      14512627
COMMENT     Contact: Kirkness EF
            The Institute for Genomic Research
            Department of Eukaryotic Genomics, TIGR, 9712 Medical Center Drive,
            Rockville, MD 20850, USA
            Tel: 301-838-0200
            Fax: 301-838-0208
            Email: ekirknes@tigr.org
            Class: shotgun.
FEATURES    Location/Qualifiers
     source          1..501
                     /organism="Canis familiaris"
                     /mol_type="genomic DNA"
                     /strain="Standard Poodle"
                     /db_xref="taxon:9615"
                     /clone_lib="Dog Library"
                     /note="Site 1: BstXI; Libraries were prepared from
                     peripheral blood"
ORIGIN

Query Match       68.0%;   Score 20.4;  DB 9;  Length 501;
Best Local Similarity   70.0%;   Pred. No. 3.1e+02;
Matches   21;  Conservative   3;  Mismatches   6;  Indels   0;  Gaps   0;

Qy      1  CAGGCUACGCGTAGAGCAUCATGATCCUGT 30
           ||||||   |||||||||||||||:|||:||
Db    263  CAGGCTAATGTTAGAGAATCATGATCCTTT 234
```

RESULT 7
AQ405865/c

```
LOCUS       AQ405865                521 bp    DNA              linear   GSS 13-MAR-1999
DEFINITION  HS_5043_B2_D11_T7 RPCI-11 Human Male BAC Library Homo sapiens
            genomic clone Plate=519 Col=22 Row=H, genomic survey sequence.
ACCESSION   AQ405865
VERSION     AQ405865.1  GI:4415853
KEYWORDS    GSS.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Hominidae; Homo.
REFERENCE   1 (bases 1 to 521)
AUTHORS     Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T., and
            Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
            Hood,L.
TITLE       Sequence-tagged connectors: A sequence approach to mapping and
            scanning the human genome
JOURNAL     Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
PUBMED      10449764
COMMENT     Contact: Mahairas GG, Wallace JC, Hood L
            High Throughput Sequencing Center
            University of Washington
            401 Queen Anne Avenue North, Seattle, WA 98109, USA
            Tel: (206) 616-3618
            Fax: (206) 616-3887
            Email: jwallace@u.washington.edu
```

Clones are derived from the human BAC library RPCI-11. For BAC
library availability, please contact Pieter de Jong
(pieter@dejong.med.buffalo.edu). Clones may be purchased from
BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
or from Resear h Genetics (info@resgen.com). BAC end Web Server:
http://www.htsc.washington.edu
Plate: 619 row: H column: 22
Seq primer: T7
Class: BAC ends
High quality sequence stop: 521.

FEATURES   Location/Qualifiers
     source   1..521
                    /organism="Homo sapiens"
                    /mol_type="genomic DNA"
                    /db_xref="taxon:9606"
                    /clone="Plate=619 Col=22 Row=H"
                    /sex="male"
                    /clone_lib="RPCI-11 Human Male BAC Library"
                    /note="Vector: pBACe3.6; Site_1: EcoRI; Site_2: EcoRI;
                    Male blood DNA was isolated from one randomly chosen donor
                    and partially digested with a combination of EcoRI and
                    EcoRI Methylase. Size selected DNA was cloned into the
                    pBACe3.6 vector at EcoRI sites"
ORIGIN

Query Match   68.0%; Score 20.4; DB 9; Length 521;
Best Local Similarity   73.3%; Pred. No. 3.2e+02;
Matches   22; Conservative   2; Mismatches   6; Indels   0; Gaps   0;

Qy   1 CAGGCUACGCGTAGAGCAUCAUCGATCCUGT 30
              |||||| |||| ||| |||||||||||||:|||
Db   209 CAGCCCACCCGTGGAGCATCATCGCTCCTTT 180

RESULT 8
AJ745356/c
LOCUS   AJ745356   631 bp   mRNA   linear   EST 01-JUL-2004
DEFINITION AJ745356 Triticum turgidum etiolated seedling 20 days old Triticum
            turgidum cDNA clone 05716R, mRNA sequence.
ACCESSION   AJ745356
VERSION   AJ745356.1 GI:49604030
KEYWORDS   EST.
SOURCE   Triticum turgidum
  ORGANISM Triticum turgidum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Pooideae; Triticeae; Triticum.
REFERENCE   1 (bases 1 to 631)
  AUTHORS   Cifarelli,R.A., D'Onofrio,O. and Lauria,G.
  TITLE   Expressed Sequences Tags (ESTs) library from totipotent cDNA of
            durum wheat
  JOURNAL   Unpublished (2003)
COMMENT   Contact: Cifarelli RA
            Biotechnology
            Metapontum Agrobios
            S.S. Jonica 106 km 448.2, 75010 Metaponto (MT), Italy.
FEATURES   Location/Qualifiers
     source   1..631
                    /organism="Triticum turgidum"
                    /mol_type="mRNA"
                    /db_xref="taxon:4571"
                    /clone="05716R"
                    /tissue_type="etiolated seedling"
                    /dev_stage="20 days old"
                    /clone_lib="Triticum turgidum etiolated seedling 20 days
                    old"
ORIGIN

Query Match   68.0%; Score 20.4; DB 1; Length 631;
Best Local Similarity   76.7%; Pred. No. 3.2e+02;
Matches   23; Conservative   1; Mismatches   6; Indels   0; Gaps   0;

Qy   1 CAGGCUACGCGTAGAGCAUCAUCGATCCUGT 30

                    |||||| |||||| |||||| ||| |||||||:|
Db   388 CAGGCCACCGCGAAGAGCTTGCTTAATCCTGT 359

RESULT 9
CA058120/c
LOCUS   CA058120   639 bp   mRNA   linear   EST 04-MAR-2003
DEFINITION ssalrga511368 mixed_tissue Salmo salar cDNA, mRNA sequence.
ACCESSION   CA058120
VERSION   CA058120.1 GI:24388363
KEYWORDS   EST.
SOURCE   Salmo salar (Atlantic salmon)
  ORGANISM Salmo salar
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei;
            Protacanthopterygii; Salmoniformes; Salmonidae; Salmo.
REFERENCE   1 (bases 1 to 639)
  AUTHORS   Rise,M.L., von Schalburg,K.R., Brown,G.D., Mawer,M.A., Devlin,R.H.,
            Kuipers,N., Busby,M., Beez-Sargent,M., Alberto,R., Gibbs,A.R.,
            Hunt,P., Shukin,R., Zeznik,J.A., Nelson,C., Jones,S.R.,
            Smailus,D.E., Jones,S.J., Schein,J.E., Marra,M.A.,
            Butterfield,Y.S., Stott,J.M., Ng,S.H., Davidson,W.S. and Koop,B.P.
  TITLE   Development and application of a salmonid EST database and cDNA
            microarray: data mining and interspecific hybridization
            characteristics
  JOURNAL   Genome Res. 14 (3), 478-490 (2004)
   PUBMED   14962987
COMMENT   Contact: Koop BF
            Centre for Biomedical Research
            University of Victoria
            PO Box 3020 STN CSC, Victoria BC, V8W 3N5, Canada
            Tel: 250 472 4067
            Fax: 250 472 4075
            Email: bkoop@uvic.ca
            Genome Sciences Centre, BC Cancer Agency
            cDNA preparation, sequencing and bioinformatics:
            Y Butterfield, R Kirkpatrick, J Asano, N Girn, R Guin, D Lee,
            S Lee, T Olson, P Pandoh, A Prahbu, D Smailus, L Spence, J Stott,
            S Taylor, G Yang, J Schein, S Jones and M Marra.
FEATURES   Location/Qualifiers
     source   1..639
                    /organism="Salmo salar"
                    /mol_type="mRNA"
                    /strain="McConnell"
                    /db_xref="taxon:8030"
                    /clone_lib="mixed_tissue"
                    /note="Vector: pCMVsport6; Library Creator: Research
                    Genetics ; Atlantic salmon tissue contributors: Carlo
                    Biagi, Mitch Uh and Robert Devlin (DFO, Vancouver, B.C.),
                    Simon Jones (PBS, Nanaimo, B.C.), Seaspring Hatchery
                    (Crofton, B.C.), Rachel Roper (University of Victoria)"
ORIGIN

Query Match   68.0%; Score 20.4; DB 5; Length 639;
Best Local Similarity   70.0%; Pred. No. 3.3e+02;
Matches   21; Conservative   3; Mismatches   6; Indels   0; Gaps   0;

Qy   1 CAGGCUACGCGTAGAGCAUCAUCGATCCUGT 30
              ||||:|| |||||| ||||||:|||| ||||:||
Db   412 CTGGCTACAAGAAGGCATCATCGCCCCTGT 383

RESULT 10
CC494201/c
LOCUS   CC494201   759 bp   DNA   linear   GSS 17-JUN-2003
DEFINITION CH240_329K8.T7 CHORI-240 Bos taurus genomic clone CH240_329K8,
            genomic survey sequence.
ACCESSION   CC494201
VERSION   CC494201.1 GI:31806325
KEYWORDS   GSS.
SOURCE   Bos taurus (cow)
  ORGANISM Bos taurus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

REFERENCE   Mammalia; Eutheria; Laurasiatheria; Cetartiodactyla; Ruminantia;
             Pecora; Bovidae; Bovinae; Bos.
AUTHORS   1 (bases 1 to 759)
             Holt.R., Stott,J., Yang,G., Barber,S., Smailus,D., Prabhu,A.-L.,
             Tsai,M., Cloutier,A., Lee,D., Girn,N., Olson,T., Mayo,M., Chan,A., Chiu,R.,
             Butterfield,Y., Kirkpatrick,R., Liu,J., Guin,R., Chan,A., Chiu,R.,
             Mathewson.C., Wye.N., Masson.A., Brown-John,M., Jones,S.,
             Schein,J., Marra,M., de Jong,P., McWilliam,S., Barris,W.,
             Dalrymple,B.P. and Tellan,R.
TITLE   Bovine BAC End Sequences from Library CHORI-240, PLATES 294 to 398
JOURNAL   Unpublished (2003)
COMMENT   Other GSSs: CH240_329K8.TARBAC13P2
             Contact: Rob Holt
             Sequencing
             The British Columbia Cancer Agency Genome Science Centre
             600 W. 10th Ave, Vancouver, British Columbia, Canada V5Z 4E6
             Tel: 604-877-6085
             Fax: 604-877-6276
             Email: rholt@bcgsc.ca
             Clones are derived from the bovine BAC library CHORI-240
             (http://www.chori.org/bacpac/bovine240.htm). For BAC library
             availability, please contact Pieter de Jong (pdejong@mail.cho.org).
             Clones may be purchased from BACPAC Resources
             (http://www.chori.org/bacpac/ordering information.htm). This work
             was undertaken as part of the International Bovine BAC Mapping
             Consortium (IBBMC) by CSIRO Livestock Industries, Australia and the
             British Columbia Genome Sciences Centre, Canada.
             Plate: 329 row: K column: 8
             Seq primer: T7
             Class: BAC ends.
FEATURES   Location/Qualifiers
     source   1..759
                     /organism="Bos taurus"
                     /mol_type="genomic DNA"
                     /strain="breed: Hereford"
                     /db_xref="taxon:9913"
                     /clone="CH240_329K8"
                     /sex="Male"
                     /cell_type="Blood"
                     /clone_lib="CHORI-240"
                     /note="Vector: pTARBAC1.3; Site_1: MboI; Site_2: MboI;
                     Hereford bull Ll Domino 99375; CHORI-240 Bovine BAC
                     library (Male) produced by Pieter de Jong"

ORIGIN
Query Match   68.0%;   Score 20.4;   DB 9;   Length 759;
Best Local Similarity   73.3%;   Pred. No. 3.3e+02;
Matches   22;   Conservative   2;   Mismatches   6;   Indels   0;   Gaps   0;

Qy   1 CAGGCUACGCGTAGAGCAUCATGATCCUGT 30
           |||||:| ||||||| |||||||||||||:|
Db   94 CAGGCTAATCAAAGAGCACCATGATCCTAT 65

RESULT 11
DN711869
LOCUS   DN711869   1332 bp   mRNA   linear EST 30-MAR-2005
DEFINITION CLJ97-H05.yld-s SHGC-CLJ Gasterosteus aculeatus cDNA clone
            CLJ97-H05 5', mRNA sequence.
ACCESSION   DN711869
VERSION   DN711869.1   GI:62076908
KEYWORDS   EST.
SOURCE   Gasterosteus aculeatus (three spined stickleback)
  ORGANISM Gasterosteus aculeatus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Gasterosteiformes;
            Gasterosteidae; Gasterosteus.
REFERENCE   1 (bases 1 to 1332)
  AUTHORS   Kingsley,D.M., Peichel,C., Balabahdra,S., Grimwood,J., Dickson,M.,
            Schmutz,J. and Myers,R.M.
  TITLE   Expressed sequence tags from Gasterosteus aculeatus
  JOURNAL   Unpublished (2003)

COMMENT   Contact: Grimwood, Jane
             Stanford Human Genome Center
             Stanford University School of Medicine
             975 S California Ave, Palo Alto, CA 94304, USA
             Tel: 650 320 5917
             Fax: 650 320 5801
             Email: jane@shgc.stanford.edu
             Plate: 97
             High quality sequence stop: 872.
FEATURES   Location/Qualifiers
     source   1..1332
                     /organism="Gasterosteus aculeatus"
                     /mol_type="mRNA"
                     /strain="Bitrufjordur marine sticklebacks, Iceland"
                     /db_xref="taxon:69293"
                     /clone="CLJ97-H05"
                     /sex="mixed male and female"
                     /tissue_type="whole larva"
                     /dev_stage="21 day old larvae collected at Swarup Stage 30
                     (J. Embryol. Exp. Morphol 6: 373-383,1958)"
                     /lab_host="DH10B (T1 phage resistant)"
                     /clone_lib="SHGC-CLJ"
                     /note="Vector: Express 1; Total and poly A+ RNA was
                     isolated from the indicated ordering tissue, and a cDNA
                     library was constructed in the Express 1 plasmid vector by
                     Open Biosystems. First strand cDNA synthesis was primed
                     with an 54 bp linker primer containing an oligodT sequence
                     preceeded by a synthetic NotI site (first strand primer:
                     5'-GACTAGTTCTAGATCGCGAGCGGCGGCGCC(T)25-3'). Following
                     second strand synthesis, cDNAs were made blunt at the end
                     corresponding to the original 5 prime end of mRNA, and
                     cloned directionally into the NotI and EcoRV sites of
                     Express 1. Note that the EcoRV site is typically destroyed
                     in the blunt end cloning, leaving a junction of the form
                     'xxxATC'(where is ATC is the second half of the EcoRV
                     site, and xxx is derived from the cDNA sequence). A map of
                     the Express 1 vector is available at:
                     http://www.openbiosystems.com/cdna_library_construction_fa
                     q.php#8 The primary library was transformed and amplified
                     in DH10B (T1 phage resistant) bacteria. Clones available
                     from Open Biosystems:
                     http://www.openbiosystems.com/stickleback"

ORIGIN
Query Match   68.0%;   Score 20.4;   DB 8;   Length 1332;
Best Local Similarity   70.0%;   Pred. No. 3.6e+02;
Matches   21;   Conservative   3;   Mismatches   6;   Indels   0;   Gaps   0;

Qy   1 CAGGCUACGCGTAGAGCAUCATGATCCUGT 30
           |||||:| ||||||| |||||||||||||:|
Db   811 CGGGCTACAAGCAGCATCATGCGCCTGT 840

RESULT 12
DN686483
LOCUS   DN686483   1351 bp   mRNA   linear EST 30-MAR-2005
DEFINITION CGX47-C07.yld-s SHGC-CGX Gasterosteus aculeatus cDNA clone
            CGX47-C07 5', mRNA sequence.
ACCESSION   DN686483
VERSION   DN686483.1   GI:62035850
KEYWORDS   EST.
SOURCE   Gasterosteus aculeatus (three spined stickleback)
  ORGANISM Gasterosteus aculeatus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Gasterosteiformes;
            Gasterosteidae; Gasterosteus.
REFERENCE   1 (bases 1 to 1351)
  AUTHORS   Kingsley,D.M., Peichel,C., Balabahdra,S., Grimwood,J., Dickson,M.,
            Schmutz,J. and Myers,R.M.
  TITLE   Expressed sequence tags from Gasterosteus aculeatus
  JOURNAL   Unpublished (2003)
  COMMENT   Contact: Grimwood, Jane

Stanford Human Genome Center
Stanford University School of Medicine
975 S California Ave, Palo Alto, CA 94304, USA
Tel: 650 320 5917
Fax: 650 320 5801
Email: jane@shgc.stanford.edu
Plate: 47
High quality sequence stop: 834.

FEATURES   Location/Qualifiers
     source   1..1351
                  /organism="Gasterosteus aculeatus"
                  /mol_type="mRNA"
                  /strain="Conner Creek sticklebacks, WA USA"
                  /db_xref="taxon:69293"
                  /clone="CGX47-C07"
                  /sex="mixed male and female"
                  /tissue_type="eyes"
                  /dev_stage="adult"
                  /lab_host="DH10B (T1 phage resistant)"
                  /clone_lib="SHGC-CGX"
                  /note="Vector: Express 1; Total and poly A+ RNA was
                  isolated from the indicated stickleback tissue, and a cDNA
                  library was constructed in the Express 1 plasmid vector by
                  Open Biosystems. First strand cDNA synthesis was primed
                  with an 54 bp linker primer containing an oligodT sequence
                  preceeded by a synthetic NotI site (first strand primer:
                  5'-GACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3'. Following
                  second strand synthesis, cDNAs were made blunt at the end
                  corresponding to the original 5 prime end of mRNA, and
                  cloned directionally into the NotI and EcoRV sites of
                  Express 1. Note that the EcoRV site is typically destroyed
                  in the blunt end cloning, leaving a junction of the form
                  'xxxATC' (where is ATC is the second half of the EcoRV
                  site, and xxx is derived from the cDNA sequence). A map of
                  the Express 1 vector is available at:
                  http://www.openbiosystems.com/cdna_library_construction_fa
                  q.php#8 The primary library was transformed and amplified
                  in DH10B (T1 phage resistant) bacteria. Clones available
                  from Open Biosystems:
                  http://www.openbiosystems.com/stickleback"

ORIGIN

Query Match   68.0%;   Score 20.4; DB 8; Length 1351;
Best Local Similarity 70.0%;   Pred. No. 3.6e+02;
Matches   21; Conservative   3; Mismatches   6; Indels   0; Gaps   0;

Qy   1 CAGGCUACGCGTAGAGCAUCATGATCCCUGT 30
           |||:|| |||||||| ||:|| ||:|
Db   797 CGGGCTACAAGCAGAGCATCATGCGCCCTGT 826

RESULT 13
DN711868/c
LOCUS   DN711868   1372 bp   mRNA   linear   EST 30-MAR-2005
DEFINITION   CLJ97-H05.xld-t SHGC-CLJ Gasterosteus aculeatus cDNA clone
             CLJ97-H05 3', mRNA sequence.
ACCESSION   DN711868
VERSION   DN711868.1   GI:62076907
KEYWORDS   EST.
SOURCE   Gasterosteus aculeatus (three spined stickleback)
ORGANISM   Gasterosteus aculeatus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
             Acanthomorpha; Acanthopterygii; Percomorpha; Gasterosteiformes;
             Gasterosteidae; Gasterosteus.
REFERENCE   1 (bases 1 to 1372)
AUTHORS   Kingsley,D.M., Peichel,C., Balabahdra,S., Grimwood,J., Dickson,M.,
             Schmutz,J. and Myers,R.M.
TITLE   Expressed sequence tags from Gasterosteus aculeatus
JOURNAL   Unpublished (2003)
COMMENT   Contact: Grimwood, Jane
             Stanford Human Genome Center
             Stanford University School of Medicine

---

975 S California Ave, Palo Alto, CA 94304, USA
Tel: 650 320 5917
Fax: 650 320 5801
Email: jane@shgc.stanford.edu
Plate: 97
High quality sequence stop: 843.

FEATURES   Location/Qualifiers
     source   1..1372
                  /organism="Gasterosteus aculeatus"
                  /mol_type="mRNA"
                  /strain="Bitrufjordur marine sticklebacks, Iceland"
                  /db_xref="taxon:69293"
                  /clone="CLJ97-H05"
                  /sex="mixed male and female"
                  /tissue_type="whole larva"
                  /dev_stage="21 day old larvae collected at Swarup Stage 30
                  (J. Embryol. Exp. Morphol 6: 373-383:1958)"
                  /lab_host="DH10B (T1 phage resistant)"
                  /clone_lib="SHGC-CLJ"
                  /note="Vector: Express 1; Total and poly A+ RNA was
                  isolated from the indicated stickleback tissue, and a cDNA
                  library was constructed in the Express 1 plasmid vector by
                  Open Biosystems. First strand cDNA synthesis was primed
                  with an 54 bp linker primer containing an oligodT sequence
                  preceeded by a synthetic NotI site (first strand primer:
                  5'-GACTAGTTCTAGATCGCGAGCGGCCGCCC(T)25-3'. Following
                  second strand synthesis, cDNAs were made blunt at the end
                  corresponding to the original 5 prime end of mRNA, and
                  cloned directionally into the NotI and EcoRV sites of
                  Express 1. Note that the EcoRV site is typically destroyed
                  in the blunt end cloning, leaving a junction of the form
                  'xxxATC' (where is ATC is the second half of the EcoRV
                  site, and xxx is derived from the cDNA sequence). A map of
                  the Express 1 vector is available at:
                  http://www.openbiosystems.com/cdna_library_construction_fa
                  q.php#8 The primary library was transformed and amplified
                  in DH10B (T1 phage resistant) bacteria. Clones available
                  from Open Biosystems:
                  http://www.openbiosystems.com/stickleback"

ORIGIN

Query Match   68.0%;   Score 20.4; DB 8; Length 1372;
Best Local Similarity 70.0%;   Pred. No. 3.6e+02;
Matches   21; Conservative   3; Mismatches   6; Indels   0; Gaps   0;

Qy   1 CAGGCUACGCGTAGAGCAUCATGATCCCUGT 30
           |||:|| |||||||| ||:|| ||:|
Db   692 CGGGCTACAAGCAGAGCATCATGCGCCCTGT 663

RESULT 14
DN686482/c
LOCUS   DN686482   1413 bp   mRNA   linear   EST 30-MAR-2005
DEFINITION   CGX47-C07.x1d-t SHGC-CGX Gasterosteus aculeatus cDNA clone
             CGX47-C07 3', mRNA sequence.
ACCESSION   DN686482
VERSION   DN686482.1   GI:62035848
KEYWORDS   EST.
SOURCE   Gasterosteus aculeatus (three spined stickleback)
ORGANISM   Gasterosteus aculeatus
             Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
             Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
             Acanthomorpha; Acanthopterygii; Percomorpha; Gasterosteiformes;
             Gasterosteidae; Gasterosteus.
REFERENCE   1 (bases 1 to 1413)
AUTHORS   Kingsley,D.M., Peichel,C., Balabahdra,S., Grimwood,J., Dickson,M.,
             Schmutz,J. and Myers,R.M.
TITLE   Expressed sequence tags from Gasterosteus aculeatus
JOURNAL   Unpublished (2003)
COMMENT   Contact: Grimwood, Jane
             Stanford Human Genome Center
             Stanford University School of Medicine
             975 S California Ave, Palo Alto, CA 94304, USA

```
            Tel: 650 320 5917
            Fax: 650 320 5801
            Email: jane@shgc.stanford.edu
            Plate: 47
            High quality sequence start: 15
            High quality sequence stop: 782.
FEATURES             Location/Qualifiers
     source          1..1413
                     /organism="Gasterosteus aculeatus"
                     /mol_type="mRNA"
                     /strain="Conner Creek sticklebacks, WA USA"
                     /db_xref="taxon:69293"
                     /clone="CGX47-C07"
                     /sex="mixed male and female"
                     /tissue_type="eyes"
                     /dev_stage="adult"
                     /lab_host="DH10B (T1 phage resistant)"
                     /clone_lib="SHGC-CGX"
                     /note="Vector: Express 1; Total and poly A+ RNA was
                     isolated from the indicated stickleback tissue, and a cDNA
                     library was constructed in the Express 1 plasmid vector by
                     Open Biosystems. First strand cDNA synthesis was primed
                     with an 54 bp linker primer containing an oligodT sequence
                     preceeded by a synthetic NotI site (first strand primer:
                     5'-GACTAGTTCTAGATCGGCAGCGGCGCC(T)25-3'). Following
                     second strand synthesis, cDNAs were made blunt at the end
                     corresponding to the original 5 prime end of mRNA, and
                     cloned directionally into the NotI and EcoRV sites of
                     Express 1. Note that the EcoRV site is typically destroyed
                     in the blunt end cloning, leaving a junction of the form
                     'xxxATC'(where is ATC is the second half of the EcoRV
                     site, and xxx is derived from the cDNA sequence). A map of
                     the Express 1 vector is available at:
                     http://www.openbiosystems.com/cdna_library_construction_fa
                     q.php#8 The primary library was transformed and amplified
                     in DH10B (T1 phage resistant) bacteria. Clones available
                     from Open Biosystems:
                     http://www.openbiosystems.com/stickleback"

ORIGIN

Query Match         68.0%;  Score 20.4;  DB 8;  Length 1413;
Best Local Similarity  70.0%;  Pred. No. 3.6e+02;
Matches   21;  Conservative   3;  Mismatches   6;  Indels   0;  Gaps   0;

Qy    1 CAGGCUACGCGTAGAGCAUCATGATCCUGT 30
         |||||:||||||||||||||||:||| ||:|
Db  673 CGGGCTACAAGCAGCAGCATCATGCGCCTGT 644

RESULT 15
CL640951
LOCUS       CL640951  CH213-1N08.T7 CH213  1669 bp  DNA  linear  GSS  06-JUL-2004
DEFINITION  CH213-1N08.T7 CH213 Gasterosteus aculeatus genomic clone CH213-1N08
            5', genomic survey sequence.
ACCESSION   CL640951
VERSION     CL640951.1  GI:49660375
KEYWORDS    GSS.
SOURCE      Gasterosteus aculeatus (three spined stickleback)
ORGANISM    Gasterosteus aculeatus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Gasterosteiformes;
            Gasterosteidae; Gasterosteus.
REFERENCE   1 (bases 1 to 1669)
  AUTHORS   Kingsley,D., Grimwood,J., Dickson,M., Schmutz,J. and Myers,R.M.
  TITLE     Expressed sequence tags from Gasterosteus aculeatus
  JOURNAL   Unpublished (2004)
COMMENT     Contact: Grimwood, Jane
            Stanford Human Genome Center
            Stanford University School of Medicine
            975 S California Avenue, Palo Alto, CA 94304, USA
            Tel: 650 320 5917
            Fax: 650 320 5801
```

```
            Email: jane@shgc.stanford.edu
            Plate: 1
            Class: BAC ends
            High quality sequence start: 35
            High quality sequence stop: 462.
FEATURES             Location/Qualifiers
     source          1..1669
                     /organism="Gasterosteus aculeatus"
                     /mol_type="genomic DNA"
                     /strain="Salmon River"
                     /db_xref="taxon:69293"
                     /clone="CH213-1N08"
                     /sex="Mixed"
                     /cell_type="Blood"
                     /clone_lib="CH213"
                     /note="Vector: pTARBAC2.1; Site_1: EcoRI; The sequence of
                     the clone was established as a mapping and sequencing
                     collaboration at the Stanford Genome Evolution Center,
                     funded by the NIH Centers of Excellence in Genomic Science
                     (CEGS) initiative (http://cegs.stanford.edu). The clone
                     was isolated from the BAC library CHORI-213 built by
                     Pieter deJong in collaboration with the Stanford Genome
                     Evolution Center (http://www.chori.org/bacpac/). Clones
                     may be purchased from BACPAC Resources
                     (http://www.chori.org/bacpac/ordering_information.html)."

ORIGIN

Query Match         68.0%;  Score 20.4;  DB 10;  Length 1669;
Best Local Similarity  70.0%;  Pred. No. 3.7e+02;
Matches   21;  Conservative   3;  Mismatches   6;  Indels   0;  Gaps   0;

Qy    1 CAGGCUACGCGTAGAGCAUCATGATCCUGT 30
         |||||:||||||||||||||||:||| ||:|
Db  493 CGGGCTACAAGCAGCAGCATCATGCGCCTGT 522

Search completed: January  8, 2006, 17:06:47
Job time : 2521 secs
```

THIS PAGE BLANK (USPTO)